

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:01:44 : Search time 16.41 Seconds
(without alignments)
632.398 Million cell updates/sec

Title: US-09-673-395a-238

Perfect score: 564
Sequence: 1 CHARLNTDSRLAKMLAYL.....NGLMHTYDSIWCNKSXN 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR.71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195.5	34.7	95	1	prostatic steroid-
2	195.5	34.7	95	1	hypothetical prote
3	195.5	34.7	95	1	D-alanyl-D-alanine
4	195.5	34.7	95	1	unknown protein, 3
5	195.5	34.7	95	1	prostatic steroid-
6	195.5	34.7	95	1	hypothetical prote
7	195.5	34.7	95	1	pristinamycin I sy
8	195.5	34.7	95	1	cell division prot
9	195.5	34.7	95	1	hypothetical prote
10	195.5	34.7	95	1	hypothetical prote
11	195.5	34.7	95	1	hypothetical prote
12	195.5	34.7	95	1	hypothetical prote
13	195.5	34.7	95	1	hypothetical prote
14	195.5	34.7	95	1	hypothetical prote
15	195.5	34.7	95	1	hypothetical prote
16	195.5	34.7	95	1	hypothetical prote
17	195.5	34.7	95	1	hypothetical prote
18	195.5	34.7	95	1	hypothetical prote
19	195.5	34.7	95	1	hypothetical prote
20	195.5	34.7	95	1	hypothetical prote
21	195.5	34.7	95	1	hypothetical prote
22	195.5	34.7	95	1	hypothetical prote
23	195.5	34.7	95	1	hypothetical prote
24	195.5	34.7	95	1	hypothetical prote
25	195.5	34.7	95	1	hypothetical prote
26	195.5	34.7	95	1	hypothetical prote
27	195.5	34.7	95	1	hypothetical prote
28	195.5	34.7	95	1	hypothetical prote
29	195.5	34.7	95	1	hypothetical prote

30	66.5	11.8	663	2	T37772
31	66.5	11.8	1601	2	T18800
32	66.5	11.6	162	2	T49124
33	66.5	11.6	429	2	D84554
34	66.5	11.6	490	2	T31646
35	66.5	11.6	603	1	W1ML8
36	66.5	11.5	137	2	F97763
37	66.5	11.5	269	2	T04394
38	66.5	11.5	322	2	E89057
39	66.5	11.5	335	2	T31712
40	66.5	11.5	349	2	T15082
41	66.5	11.5	874	1	S07380
42	66.5	11.5	975	2	A86258
43	66.5	11.5	1255	2	T06267
44	66.5	11.5	2145	2	S61041
45	66.5	11.4	113	2	JC4143

ALIGNMENTS

RESULT 1

prostatic steroid-binding protein chain C3 precursor - rat
B0RT3
N:Alternate names: prostatic
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C:Accession: A92395; A92433; A91108; A42392; A03250
R:Paraker, M.G.; White, R.; Hurst, H.; Needham, M.; Tilly, R.
J. Biol. Chem. 258, 12-15, 1983
J. Biol. Chem. 258, 12-15, 1983
A:Title: Prostatic steroid-binding protein. Isolation and characterization of C3 gene
A:Reference number: A92395; MUID:83082848
A:Accession: A92395
A:Molecule type: mRNA
A:Residues: 1-95 <PAR>
A:Cross-references: GB:V01263; GB:J00777; NID:956993; PIDN:CA24577.1; PID:956994
R:Vlakochill, D.H.; Perry, S.T.; Lea, O.A.; Stafford, D.W.; Wilson, E.M.; French, F.S.
J. Biol. Chem. 258, 8861-8866, 1983
A:Title: Isolation of two genomic sequences encoding the M-r = 14,000 subunit of rat
A:Reference number: A92433; MUID:83238526
A:Accession: A92433
A:Molecule type: mRNA
A:Residues: 1-78, 'S', 80-95 <VIS>
R:Peeters, B.; Rombauts, W.; Mous, J.; Heyns, W.
Eur. J. Biochem. 115, 115-121, 1981
A:Title: Structural studies on rat prostatic binding protein. The primary structure o
A:Reference number: A91108; MUID:81188769
A:Accession: A91108
A:Molecule type: protein
A:Residues: 19-95 <PEP>
R:Tan, J.A.; Marschke, K.B.; Ho, K.C.; Perry, S.T.; Wilson, E.M.; French, F.S.
J. Biol. Chem. 267, 4456-4466, 1992
A:Title: Response elements of the androgen-regulated C3 gene.
A:Reference number: A42392; MUID:92165796
A:Accession: A42392
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-16, 'T', 18-95 <TRAN>
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBI:83851, NCBI:83856)
C:Comment: C3 is encoded by two unique genes that differ from each other only in the
C:Comment: Steroid-binding protein, the principal androgen-dependent secretory protei
ng C2 and C3 chains. The chains of each dimer are linked by disulfide bonds. This pro
C:Superfamily: uteroglobin
C:Keywords: glycoprotein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-95/Product: prostatic steroid-binding protein C3 chain #status experimental <MAT
F:35/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 34.7% Score 195.5; DB 1; Length 95;
Best Local Similarity 35.8%; Pred. No. 8e-13;

```

Matches      34; Conservative      30; Mismatches      30; Indels      1; Gaps      1;
OY          14 MKLLMTMLTALLALHICVAD-SGCKLLEDVVEKTIINSIDISIEYKELLQEIFISDPAANAEM 72
             |||:: :: :: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db           1 MKLVFLFLVLTIPICGYASGGSGSLIDEVIROGTINSTVTLDHYMKLVKVPYODHFTEKAV 60
OY          73 GKPKOCFLNOSHRILTAKNFGLMHTIYDSITWCNMKS 107
             :|||||:: ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
Db           61 KFKOCFLDQTDKTLENVGVMMAEAIINSSSCOPPS 95
RESULT      2
T38449
C:Biochemical protein SPAC27E2.02 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Cdate: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38449
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21749
A:Accession: T38449
A:status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-280 <NMR>
A:Cross-references: EMBL:Z98978; PIDN:CA811676.1; GSPDB:GN00066; SPDB:SPAC27E2.02
A:Experimental source: strain 972n-; cosmid c27E2
C:Genetics:
A:Gene: SPDB:SPAC27E2.02
A:Map position: 1
A:Introns: 185/3; 225/3

```

```

Query Match      13.6%   Score 76.5; DB 2; Length 280;
Best Local Similarity 30.3%; Pred. No.3,1;
Matches 23; Conservative 10; Mismatches 28; Indels 15; Gaps 3;

OY    37 LLEDMVETKINSDISPEYKELLOEPIFSD---AAAFAMGKFKQ-----CFILNQSHRT 86
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB    68 LAEDVILLSVATGDVCITFTSMDLKELVLDIAEQAAAARESKSQEESDKETPYMLNKRHYV 127

OY    87 LKNFGIMMHTVYDSIW 102
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
DB    128 AKT-----PEIQDEPW 138

RESULT      3
H87530
C:d-alanyl-D-alanine dipeptidase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: H87530
R:Nierman, W.C.; Feldbljum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; NCID:21173698; PMID:11259647
A:Accession: H87530
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-212 <STO>
A:Cross-references: GB:AE005673; NID:g13423786; PIDN:AKK24244.1; GSPDB:GN00148
C:Genetics:
C:Gene: CC2273
C:Superfamily: Enterococcus faecium transposon Tn1546 D-alanine-D-alanine dipeptidase

Query Match      13.5%   Score 76; DB 2; Length 212;
Best Local Similarity 29.9%; Pred. No.2,6;
Matches 26; Conservative 13; Mismatches 16; Indels 32; Gaps 5;

OY    31 ADSGCKLLEDVEKT-----INSDISPEYK--LLDFIFDSDAAAFAMG 73
      | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

```

Db      19  AETGKTLASLDVDTYTTTTPGICKIDIRYAGANNFNNGPILEYKSAAYLQR-----PAAEALG 73
OY      74  KFKOCFLNQSHRTL--KNEGLMAHTVY 98
           || | | | | | | | | | |
Db      74  RI-----HRAALAKGYGLLIHDAY 92

RESULT  4
A96656
unknown protein, 38394-36551 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A96656
C:Organism: A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizart, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.R.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzita
A:Authors: Salzbberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yi, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference Number: A86141; MUID:21016719
A:Accession: A96656
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-590 <STO>
A:Cross-References: GB:AE005173; NID:q10092251; PIDN:AMG12665.1; GSPDB:GN00141
C:Genetics:
A:gene: pt6m19.15
A:Map position: 1

```

```

Query Match      13.0% Score 73.5; DB 2; Length 590;
Best Local Similarity 32.38; Pred. NO.14;
Matches 21; Conservative 8; Mismatches 23; Indels 13; Gaps 3;

QY      26 LHHGADSG-----CKLLEDVYEKTINSIDISPEKELLOEFIDSDAAEA-----NGK 74
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      291 LISLCATGRMSDASRLSLDMLKEKNIPDLVF--FNALIDAVRKGLVEAKELYDEMYK 348

DB      349 SKHCF 353

RESULT    5
BORT2
prostatic steroid-binding protein chain C2 precursor - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 24-Sep-1999
C.Accession: A03251; A26671
R.Parker, M.; Needham, M.; White, R.
Nature 298, 92-94, 1982
A.Title: Prostatic steroid binding protein: gene duplication and steroid binding.
A.Reference number: A93286; MUID:82220075
A.Accession: A03251
A.Molecule type: mRNA
A.Residues: 1-98 <PAR>
A.Cross-references: GB:J00776; NID:g206448; PIDN:AAA51641.j; PID:g206450
R.Delaey, B.; Dirckx, L.; Decourt, J.L.; Claessens, F.; Peeters, B.; Rombaux, W.
Nucleic Acids Res. 15, 1627-1641, 1987
A.Title: Rat prostatic binding protein: the complete sequence of the C2 gene and its
A.Reference number: A26671; MUID:87146484
A.Accession: A26671
A.Molecule type: DNA
A.Residues: 1-23, 'Q', 26-86, 'I', 88-94, 'W/LQINFPGRGFSEIN' <DEL>
A.Cross-references: GB:X05034; NID:g56857; PIDN:CAA28708.1; PID:g56858
C.Comment: Steroid-binding protein, the principal secretory protein in rat prostatic
The chains of each dimer are linked by disulfide bonds.
C:Superfamily: uteroglobin
C:Keywords: heterotetramer; prostate; steroid binding

```


Query Match 12.4%; Score 70; DB 2; Length 303;
 Best Local Similarity 35.2%; Pred. No. 15;
 Matches 19; Conservative 10; Mismatches 17; Indels 8; Gaps 1;

OY 38 LEDVKTINSDISIPEYKELDF-----IDSDAAEAMGKFKOCFLNOS 83
 DB 37 LNELELLISSDISIVTHIEEFKNVFKDKTIDSDYKEALAKLEOOLSKS 90

RESULT 10

hypothetical protein C39E6.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Apr-2000

C:Accession: T29741

R:Feature: B. Le, T.

submitted to the EMBL Data Library, February 1996

A:Description: The sequence of C. elegans cosmid C39E6.

A:Reference number: 220676

A:Accession: T29741

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Residues: 1-457 <FOL>

A:Cross-references: EMBL:U49944; PIDN:AAA93419.1; CESP:C39E6.6

C:Genetics:

A:Gene: CESP:C39E6.6

A:Introns: 58/3; 107/1; 169/3; 253/3; 274/2; 329/2; 366/2; 438/2

C:Supportfamily: neurokinin 1 receptor

Query Match 12.4%; Score 70; DB 2; Length 457;
 Best Local Similarity 28.3%; Pred. No. 24;
 Matches 30; Conservative 16; Mismatches 32; Indels 28; Gaps 6;

OY 1 CHARLNTSSRIAMKLIVMLA-----ALLHCVADSGCKL-----LEDVKT-- 45
 DB 178 CTEKESAKSRA--YTMIVMLAQFVPFAVAFYANIVSVLSKRAQTKRKVERISA 235

OY 46 INSDISIP-----EYKELDFIDSDAAEAMGKFKOCFLNOSHRT 86
 DB 236 LSSSCAFPSHGILQYENELNFKDF-----KEKORVVLQNKRRT 275

RESULT 11

hypothetical protein T12K4.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47321

R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Rudd, S.; Lemcke, K.; May

submitted to the Protein Sequence Database, April 2000

A:Reference number: 224460

A:Accession: T47321

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-412 <MON>

A:Cross-references: EMBL:AL138640

A:Experimental source: cultivar Columbia; BAC clone T12K4

C:Genetics:

A:Map position: 3

A:Introns: 30/3

A:Note: T12K4.80

Query Match 12.3%; Score 69.5; DB 2; Length 412;
 Best Local Similarity 27.5%; Pred. No. 24;
 Matches 25; Conservative 11; Mismatches 40; Indels 15; Gaps 3;

OY 26 LLHCYADSGCKL---EDVKTINSDISIPEYKELDFID---SDAAEAMGKFKOC 79
 DB 15 LLSCTYTSSAPSIKKSKSLRDSOKIQIVDAPLVLQTLQRRLPDVAHEIFLQTKSVN 74

OY 80 LNOQHRTL-----KNFGLMMHTVYDSI 101

DB 75 LLPVRTLCAIMLCFAENGFLVRLARRITMDEI 105

RESULT 12

protein F1N19.15 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96669

R:Theologos, A.; Becker, J.R.; Palm, G.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hutzler, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, K.; Matzida

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salberg, S.L.; Scharitz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venier, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A6141; MUID:21016719

A:Accession: C96669

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1048 <STO>

A:Cross-references: GB:A6005173; NID:g6633829; PIDN:AAF19688.1; GSPDB:GN00141

C:Genetics:

A:Gene: F1N19.15

A:Map position: 1

Query Match 12.3%; Score 69.5; DB 2; Length 1048;
 Best Local Similarity 23.4%; Pred. No. 66;
 Matches 22; Conservative 20; Mismatches 15; Indels 37; Gaps 6;

OY 29 CYA---DSGCKLEDVKTINSDI-----SIPE-----KELDFIDSDA 67
 DB 222 CAGSRNSDAARFMRDRMKRSINPDVFTALIDVFKQGNLDEOELKEMIOSSVDPN- 280

OY 68 AAEAMGKFKOCFLNOSHRTLNKFLMMH--TVYDS 100
 DB 281 -----NVTYNSIIN-GLCMHGRITDA 300

RESULT 13

hypothetical protein AF1487 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: F69435

R:Klenk, H.P.; Clayton, K.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod

der, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirschner, E.

Glodek, A.; Zhou, L.; Overbeek, R.; Goceyne, J.D.; Weidman, J.F.; McDonald, H.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Colton, M.D.; Springs, T.; Arltach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venier, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch

A:Reference number: A69250; MUID:98049343

A:Accession: F69435

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-219 <KLE>

A:Cross-references: GB:A6001000; GB:A600782; NID:g2689323; PIDN:AA89768.1; PID:g264

Query Match 12.2%; Score 69; DB 2; Length 219;
 Best Local Similarity 24.5%; Pred. No. 14;
 Matches 25; Conservative 18; Mismatches 39; Indels 20; Gaps 5;

OY 7 TDSSRLAKLMLVLMALLLHCYADSGCKLEDVKTINSDISIPEYKELDF----- 61
 DB 2 SDMRL-MILFLVGLAVL-----SGCATLS--VSKYNNKGSVSSTLYLVINTSPFY 52

OY 62 -FIDSDAAAEAMGKFCFLNOSHRTLNFGILMHTVYDSIW 102
 DB 53 GLLAEGAKKEGYESLRESFLSEIPREMRD-----KVSXDEW 89

RESULT 14

T47569
 hypothetical protein F24B22.60 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #lexl_change 20-Apr-2000
 C:Accession: T47569
 R:Boecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Ouelier, F.; Salanoubat, M.
 submitted to the Protein Sequence Database, January 2000
 A:Reference number: #23016
 A:Accession: T47569
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-638 <HIO>
 A:Cross-references: EMBL:AL132957
 A:Experimental source: cultivar Columbia; BAC clone F24B22
 C:Genetics:
 A:Map position: 3
 A:Introns: 192/3; 217/1; 240/3; 270/2; 314/3; 392/2; 551/2
 A:Note: F24B22.60

Query Match 12.2%; Score 69; DB 2; Length 638;
 Best Local Similarity 21.7%; Pred. No. 43;
 Matches 23; Conservative 22; Mismatches 33; Indels 28; Gaps 3;

OY 4 RLNTDSSRLAMKILMVLMLAALLHCYADSG-----CKLEPDMVEKTIINSDISIPRYKE 57
 DB 116 RSGRNVGRRLTGLMLLVASVFLRYLMGVGVVDHARLKEFVVVTRLDDMSMAO--- 172
 OY 58 LLOEFIDSDAAAEAM-----GKFCFLNOSHRT 86
 DB 173 --RVAENQASQPMRYLEKPIPEIWMQPSGNTKOCVTRPKNYT 216

RESULT 15

T01889
 hypothetical protein F8M12.4 - Arabidopsis thaliana
 N:Alternate names: hypothetical protein F25124.80
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #lexl_change 21-Jul-2000
 C:Accession: T01889; T04282
 R:Madson, C.; Graves, T.; Colton, M.; Modde, T.
 submitted to the EMBL data library, July 1998
 A:Description: The sequence of A. thaliana F8M12.
 A:Reference number: Z14450
 A:Accession: T01889
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-173 <MAD>
 A:Cross-references: EMBL:AF080118; NID:q3513725; PID:q3513732
 A:Experimental source: cultivar Columbia
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, A.
 submitted to the Protein Sequence Database, March 1999
 A:Reference number: Z15261
 A:Accession: T04282
 A:Molecule type: DNA
 A:Residues: 1-173 <BEV>
 A:Cross-references: EMBL:AL049525
 A:Experimental source: cultivar Columbia; BAC clone F25124
 C:Genetics:
 A:Map position: 4
 A:Note: F8M12.4; F25124.80
 C:Superfamily: Arabidopsis thaliana hypothetical protein F8M12.2

Query Match 12.1%; Score 68.5; DB 2; Length 173;
 Best Local Similarity 24.3%; Pred. No. 12;
 Matches 27; Conservative 17; Mismatches 36; Indels 31; Gaps 6;

OY 24 ALLHGYADSGC-----KLEDMV-----EKTINSDISIPRYKEILQEFJDSIDAAK 70
 DB 2 APLLHRYSNIGCGKRYTVRFCDKTIYDTNDESRITNEELNSGPTIKIYRDPDEDQGYL 61
 OY 71 AMGKFCFLNOSHRTLNFGILMHTVYDSIW 105
 DB 62 LNYDGYMLMINDNG--KNKGIDYDLVHWTAVYDDINIDIVYDEMKIWCEI 110

Search completed: June 20, 2002, 11:03:07
 Job time: 83 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:02:44 ; Search time 11.86 Seconds
(Without alignments)
352.589 Million cell updates/sec

Title: US-09-673-395A-238

Perfect score: 564

Sequence: 1 CHARLNTDSSRLAMKRLNVL.....NEGLMHTVDSIMCKMKN 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	496	87.9	95	1	MGBB_HUMAN
2	281	49.8	93	1	MGBA_HUMAN
3	195.5	34.7	95	1	PSC3_RAT
4	71.5	12.7	131	1	YRS9_MYCTU
5	70	12.4	303	1	PTSY_RICPR
6	69	12.2	112	1	PSC2_RAT
7	68.5	12.1	684	1	TC10_YEAST
8	68	12.1	1071	1	PR16_YEAST
9	67.5	12.0	268	1	TRPC_ACTCA
10	67.5	12.0	871	1	SC10_YEAST
11	66.5	11.8	234	1	RADC_HAEIN
12	66.5	11.8	663	1	TAZ1_SCHPO
13	65.5	11.6	162	1	IL15_MOUSE
14	65.5	11.6	603	1	VEL_HPV08
15	65	11.5	2144	1	G1T1_YEAST
16	64.5	11.4	113	1	MIR_CAI
17	64.5	11.4	153	1	Y156_UREPA
18	64.5	11.4	162	1	VE1_HPV21
19	64.5	11.4	603	1	VE1_HPV21
20	64.5	11.4	611	1	DYNL_ARATH
21	64	11.3	473	1	MAIR_YEAST
22	64	11.3	708	1	TRAB_HUMAN
23	63.5	11.3	681	1	CAO2_RABIT
24	63.5	11.3	1294	1	YAB3_SCHPO
25	63	11.2	264	1	PSA4_DROME
26	63	11.2	1022	1	SCA4_RICPR
27	62.5	11.1	91	1	UTER_RABIT
28	62.5	11.1	223	1	YMA2_YEAST
29	62.5	11.1	610	1	VG12_BPT4
30	62	11.0	111	1	PSC1_RAT
31	62	11.0	938	1	V120_HSV7J
32	61.5	10.9	570	1	HEM1_KTULA
33	61.5	10.9	734	1	METE_THIEMA

34	61.5	10.9	2190	1	CCAD_CHICK	073700 gallus gall
35	61	10.8	63	1	YEBW_ECOLI	P76275 escherichia
36	61	10.8	142	1	GRPA_VACCC	P20494 vaccinia vi
37	61	10.8	605	1	VE1_HPV14	P36721 human papil
38	61	10.8	606	1	VE1_HPV5H	P26542 human papil
39	61	10.8	732	1	K086_MOUSE	P27641 mus musculu
40	61	10.8	830	1	JIP2_MOUSE	Q9er49 mus musculu
41	61	10.8	874	1	RPOL_BPSF6	P06221 bacterioph
42	61	10.8	1272	1	Y228_METJA	O60287 methanococ
43	61	10.8	1966	1	CCAF_HUMAN	O60840 homo sapien
44	60.5	10.7	120	1	RHPA_BUCAL	P57457 buchnera ap
45	60.5	10.7	180	1	MIP2_MOUSE	P11589 mus musculu

ALIGNMENTS

RESULT	ID	MGBB_HUMAN	STANDARD	PRT	95 AA
AC	075556				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Mammaglobin B precursor (Mammaglobin 2) (Lipophilin C) (Lactogloblin).				
GN	MGB2 OR UGB3 OR LIPHC.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99026127; PubMed=9806831;				
RA	Becker R.M., Darrow C., Zimonjic D.B., Popescu N.C., Watson M.A.,				
RA	Fleming T.P.;				
RT	*Identification of mammaglobin B, a novel member of the uteroglobin				
RT	gene family.*;				
RL	Genomics 54:70-78(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99167354; PubMed=10066439;				
RA	Zhao C., Nguyen T., Yusifov T., Glasgow B.J., Lehrer R.I.;				
RA	Willcox M.D., Gooley A.A., Williams K.L., Morris G.A., Walsh B.J.;				
RT	*Establishment of the human reflex tear two-dimensional polyclacrylamide				
RT	gel electrophoresis reference map: new proteins of potential				
RT	diagnostic value.*;				
RL	Electrophoresis 18:2811-2815(1997).				
RN	[4]				
RP	SEQUENCE OF 19-46 AND 60-78, AND MASS SPECTROMETRY.				
RX	MEDLINE=98385871; PubMed=9720917;				
RA	Leher R.I., Xu G., Abduragimov A., Dinh N.N., Qu X.-D., Martin D.,				
RA	Glasgow B.J.;				
RT	*Lipophilin, a novel heterodimeric protein of human tears.*;				
RL	FEBS Lett. 432:163-167(1998).				
CC	- FUNCTION: MAY BIND ANDROGENS AND OTHER STEROIDS, MAY ALSO BIND				
CC	ESTRADIOL, A CHEMOTHERAPEUTIC AGENT USED FOR PROSTATE CANCER.				
CC	MAY BE UNDER TRANSCRIPTIONAL REGULATION OF STEROID HORMONES.				
CC	- SUBUNIT: HETERODIMER OF A LIPOPHILIN A AND A LIPOPHILIN C				
CC	(MAMMAGLOBIN B) MONOMER ASSOCIATED HEAD TO HEAD.				
CC	- TISSUE SPECIFICITY: EXPRESSED IN THYMUS, TRACHEA, KIDNEY, STEROID				
CC	RESPONSIVE TISSUES (PROSTATE, UTERUS, BREAST AND OVARY),				
CC	AND SALIVARY GLAND.				
CC	- MASS SPECTROMETRY: MW=854.94; METHOD=Electrospray, RANGE=19-95.				
CC	- SIMILARITY: BELONGS TO THE UTEROGLIBIN FAMILY. LIPOPHILIN				
CC	SUBFAMILY.				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF071219; AAC79996.1; -
 CC EMBL: AJ224173; CA11865.1; -
 CC MIM: 604398; -
 DR InterPro: IPR003627; Mammaglobin_prostin.
 DR InterPro: IPR000329; Uterogloblin.
 DR Pfam: PF01099; Uterogloblin; 1.
 DR ProDom: PD029354; Mammaglobin_prostin; 1.
 DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE_NEG.
 DR PROSITE: PS00404; UTEROGLOBIN_2; FALSE_NEG.
 KW Signal: Glycoprotein.
 FT SIGNAL 1 18
 FT CHAIN 19 93 MAMMAGLOBIN B.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 95 AA: 10884 MW: 0719738289F8F8D CRC64;

Query Match 87.9%; Score 496; DB 1; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.2e-43;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWVLMALALHCHYADSGCKLEDMVEKTSIPYKELLOEFLDSDAAAFAMG 73
 DB 1 MKLWVLMALALHCHYADSGCKLEDMVEKTSIPYKELLOEFLDSDAAAFAMG 60

OY 74 KFKQCFLNQSHRTLNKFGIMHTVYDSIWMKSN 108
 DB 61 KFKQCFLNQSHRTLNKFGIMHTVYDSIWMKSN 95

RESULT 2

MGBA_HUMAN STANDARD: PRT; 93 AA.

AC 013396;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mammaglobin A precursor (Mammaglobin 1).
 DE MGB1 OR UGB2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=breast;
 RX MEDLINE=96223698; PubMed=8631025;
 RA Watson M.A., Fleming T.P.;
 RT "Mammaglobin, a mammary-specific member of the uterogloblin gene family, is overexpressed in human breast cancer.";
 RL Cancer Res. 56:860-865(1996).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98147311; PubMed=9488047;
 RA Watson M.A., Darlow C., Zimonjic D.B., Popescu N.C., Fleming T.;
 RT "Structure and transcriptional regulation of the human mammaglobin gene, a breast cancer associated member of the uterogloblin gene family localized to chromosome 11q13.";
 RL Oncogene 16:817-824(1998).
 CC -1- TISSUE SPECIFICITY: MAMMARY-SPECIFIC. OVER-EXPRESSED IN BREAST CANCER.
 CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: U33147; AAC50375.1; -
 CC EMBL: AF015224; AAC39608.1; -
 CC MIM: 605562; -
 DR InterPro: IPR003627; Mammaglobin_prostin.
 DR InterPro: IPR000329; uterogloblin.
 DR Pfam: PF01099; Uterogloblin; 1.
 DR ProDom: PD029354; Mammaglobin_prostin; 1.
 DR SMART: SM00096; UTRG; 1.
 DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE_NEG.
 DR PROSITE: PS00404; UTEROGLOBIN_2; 1.
 KW Signal: Glycoprotein.
 FT SIGNAL 1 18
 FT CHAIN 19 93 MAMMAGLOBIN A.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 93 AA: 10499 MW: 28968C43BF053E2 CRC64;

Query Match 49.8%; Score 281; DB 1; Length 93;
 Best Local Similarity 58.7%; Pred. No. 7.1e-22;
 Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 14 MKLWVLMALALHCHYADSGCKLEDMVEKTSIPYKELLOEFLDSDAAAFAMG 73
 DB 1 MKLWVLMALALHCHYADSGCKLEDMVEKTSIPYKELLOEFLDSDAAAFAMG 60

OY 74 KFKQCFLNQSHRTLNKFGIMHTVYDSIWMKSN 105
 DB 61 KFKQCFLNQSHRTLNKFGIMHTVYDSIWMKSN 92

RESULT 3

PSC3_RAT STANDARD: PRT; 95 AA.

AC P02780; 063463;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostatic steroid-binding protein. Isolation and characterization of peptide C3).
 DE Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.

RC MEDLINE=83082848; PubMed=6294095;
 RA Parker M.G., White R., Hurst H., Needham M., Tilly R.;
 RT "Prostatic steroid-binding protein. Isolation and characterization of C3 genes.";
 RL J. Biol. Chem. 258:12-15(1983).

RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83328526; PubMed=6190812;
 RA Viskochil D.H., Perry S.T., Lea O.A., Stafford D.W., Wilson E.M., French F.S.;
 RT "Isolation of two genomic sequences encoding the Mr = 14,000 subunit of rat prostatein.";
 RL J. Biol. Chem. 258:8861-8866(1983).

RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92165796; PubMed=1537831;
 RA Tan J.A., Marschke K.B., Ho K.-C., Perry S.T., Wilson E.M., French F.S.;
 RT "Response elements of the androgen-regulated C3 gene.";
 RL J. Biol. Chem. 267:4456-4466(1992).

[4]


```

ERRATUM
RX MEDLINE=92218467; PubMed=1339454;
RA Tan J.A., Marschke K.B., Ho K.C., Perry S.T., Wilson E.M.,
RL French F.S.,
RL J. Biol. Chem. 267:7958-7958(1992).
RN 151
RP SEQUENCE OF 19-95.
RX MEDLINE=81188769; PubMed=7014218.
RA Peeters B., Romhans W., Mous J., Heyns W.;
RT "Structural studies on rat prostatic binding protein. The primary
RT structure of its glycosylated component C3.";
RL Eur. J. Biochem. 115:115-121(1981).
CC -1- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND. STEROID-BINDING PROTEIN;
CC CAN BIND NON-POLAR STEROIDS, CHOLESTEROL AND A GROUP OF SMALL
CC PROLINE-RICH PEPTIDES.
CC -1- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
CC C1, C2 AND C3. THESE FORM COVALENT C1:C3 (P) AND C2:C3 (S)
CC HETERODIMERS WHOSE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
CC (C1:C3/C3:C2) PROSTATEIN MOLECULES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: VENTRAL PROSTATE.
CC -1- INDUCTION: ANDROGEN DEPENDENT, AS SHOWN BY THE DECREASE IN THE
CC LEVEL OF THE PROTEIN FOLLOWING CASTRATION.
CC -1- MISCELLANEOUS: C3 IS ENCODED BY TWO DIFFERENT GENES.
CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPONHILIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL: Y01263; CAA24577.1; -.
DR EMBL: M71245; AAA41965.1; -.
DR PIR: A03250; B0R73.
DR InterPro: IPR003627; Mammagln.prostin.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Mammagln.prostin; 1.
DR ProDom: PD029354; Mammagln.prostin; 1.
DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE_NRG.
DR PROSITE: PS00404; UTEROGLOBIN_2; 1.
KW Signal; Glycoprotein; Steroid-binding.
FT SIGNAL 1 18
FT CHAIN 19 95
FT PROSTATIC STEROID-BINDING PROTEIN C3
FT CHAIN
FT CARBOHYD 35 35
FT N-LINKED (GLCNAC. . .)
FT CONFLICT 53 53 D -> A (IN REF. 3).
FT CONFLICT 79 79 G -> S (IN REF. 2).
FT CONFLICT 79 79
DR SEQUENCE 95 AA: 10730 MW: F7F7ELAAC882E375 CRC64:
Query Match 34.7%; Score 195.5; DB 1; Length 95;
Best Local Similarity 35.8%; Pred. No. 3, 3e-13;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1
OY 14 MLLVAVMLAALLHGYAD-SGCKLDEDMVEKTIINDISIPYKELIDFIDSDAAAEAM 72
DB 1 MLLVAVMLAALLHGYAD-SGCKLDEDMVEKTIINDISIPYKELIDFIDSDAAAEAM 72
DB 1 MLLVAVMLAALLHGYAD-SGCKLDEDMVEKTIINDISIPYKELIDFIDSDAAAEAM 72
OY 73 GKFOCFILNSHRLTKNFGMLHGHVYSINCMKMS 107
DB 73 GKFOCFILNSHRLTKNFGMLHGHVYSINCMKMS 107
DB 73 GKFOCFILNSHRLTKNFGMLHGHVYSINCMKMS 107
OY 61 KFKOCFLDQTKLTLENGVMMHKAIFNSSCQPS 95
DB 61 KFKOCFLDQTKLTLENGVMMHKAIFNSSCQPS 95
RESULT 4
YR59_MYCTU
ID YR59_MYCTU STANDARD; PRT; 131 AA.
AC 033301;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DE DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DT Hypothetical 14.4 kDa protein P42759C.
GN RV2759C OR MT2829 OR MTV002.24C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN RP SEQUENCE FROM N.A.
RC STRAIN-H37RV:
RC MEDLINE=96295987; PubMed=9634230;
RA Cole S.T., Brosch K., Parkhill J., Garner T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.K. III, Tekle A.F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Horsby T., Jorgensen K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RC Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolony A.J., Nelson W.C., Umayam L.A., Ermolaeva M.A., Salzberg S.L.,
RA Delcher A.L., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UP00110 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL008967; CAA15555.1;
CC DR EMBL: AE007110; AMK47148.1;
CC DR TIGR: MT2829;
CC DR TubercuList: K02759c;
CC DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA; 14372 MW; 9C72AKDACG74065B CRC64;

```

GN FTSY OR RP775.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 NCBI_TaxId=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=97419517; PubMed=9274032;
 RA Andersson J.O., Andersson S.G.E.;
 RT "Genomic rearrangements during evolution of the obligate
 intracellular parasite Rickettsia prowazekii as inferred from an
 analysis of 52015 bp nucleotide sequence.";
 RL Microbiology 143:2783-2795(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sitcheritz-Ponten T., Almarik U.C.M., Podowski R.M., Naslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria.";
 RL Nature 396:133-140(1998).
 RN [3]
 RP SEQUENCE OF 168-303 FROM N.A.
 RC STRAIN=MADRID E, AND B;
 RX MEDLINE=99416441; PubMed=10486973;
 RA Andersson J.O., Andersson S.G.E.;
 RT "Genome degradation is an ongoing process in Rickettsia.";
 RL Mol. Biol. Evol. 16:1178-1191(1999).
 CC -1- FUNCTION: FUNCTIONAL HOMOLOG OF SRP RECEPTOR. PROBABLY INVOLVED IN
 CC THE RECEPTION AND INSERTION OF A SUBSET OF PROTEINS AT THE
 CC MEMBRANE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Y11784; CAA72477.1;
 DR EMBL: AJ235273; CAA15202.1;
 DR EMBL: AJ238755; CAB56084.1;
 DR EMBL: AJ238756; CAB56088.1;
 DR HSSP: P10121; ICTS.
 DR InterPro: IPR000897; SRP54.
 DR Pfam: PF00448; SRP54_N; 1.
 DR Pfam: PF02881; SRP54_N; 1.
 DR ProDom: PD000819; SRP54; 1.
 DR PROSITE: PS00300; SRP54; 1.
 KM Signal recognition particle: GTP-binding; RNA-binding; Membrane;
 KM Cell division; Complete proteome.
 FT NP_BIND 108 115 GTP (BY SIMILARITY).
 FT NP_BIND 190 194 GTP (BY SIMILARITY).
 FT NP_BIND 254 257 GTP (BY SIMILARITY).
 FT SEQUENCE 303 AA; 33313 MW; BF3FE9383B7007 CRC64;
 Query Match 12.4%; Score 70; DB 1; Length 303;
 Best Local Similarity 35.2%; Pred. No. 5.9; Mismatches 17; Indels 8; Gaps 1;
 Matches 19; Conservative 10;
 QY 38 LEDWYKTNISDISPEYKELQEF-----IDSDAAAEAMGKFCQFLNQS 83
 DB 37 LNFLELLSSDISISVTHIEEFKNVAFDKTIDSDYKKAIAKLISQOLSKS 90
 RESULT 6

PSC2_RAT
 ID PSC2_RAT STANDARD: PRT; 112 AA.
 AC P02781;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostatic steroid-binding protein C2 chain precursor (Prostatein
 DE peptide C2).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87146484; PubMed=2881277;
 RA Delaey B., Dirckx L., Decourt J.-L., Claessens F., Peeters B.,
 RA Rombauts W.;
 RT "Rat prostatic binding protein: the complete sequence of the C2 gene
 RT and its flanking regions.";
 RL Nucleic Acids Res. 15:1627-1641(1987).
 RN [2]
 RP SEQUENCE OF 21-112.
 RX MEDLINE=83209619; PubMed=6343081;
 RA Peeters B., Heyns W., Mous J., Rombauts W.;
 RT "Structural studies on rat prostatic binding protein. The primary
 RT structure of component C2 from subunit S.";
 RL Eur. J. Biochem. 132:669-679(1983).
 RN [3]
 RP SEQUENCE OF 1-100 FROM N.A.
 RX MEDLINE=82220075; PubMed=6896362;
 RA Parker M., Needham M., White R.;
 RT "Prostatic steroid binding protein: gene duplication and steroid
 RT binding.";
 RL Nature 298:92-94(1982).
 CC -1- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
 CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND.
 CC -1- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
 CC C1, C2 AND C3. THESE FORM COVALENT C1:C3 (F) AND C2:C3 (S)
 CC HETERODIMERS WHOSE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
 CC (C1:C3/C3:C2) PROSTATEIN MOLECULES.
 CC -1- SUBCELLULAR LOCATION: secreted.
 CC -1- PTM: LINKED BY THREE DISULFIDE BONDS TO C3.
 CC -1- MISCELLANEOUS: THE HETERODIMER CAN BIND NON-POLAR STEROIDS,
 CC CHOLESTEROL AND A GROUP OF SMALL PROLINE-RICH PEPTIDES.
 CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
 CC SUPERFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X05034; CAA28708.1;
 DR EMBL: V01256; CAA24569.1;
 DR EMBL: J00776; AAA51641.1;
 DR PIR: A03251; BORT2.
 DR PIR: A26671; A26671.
 DR InterPro: IPR003628; Uteroglobln-sub.
 DR InterPro: IPR000329; Uteroglobln.
 DR ProDom: PD012475; Uteroglobln-sub; 1.
 DR PROSITE: PS00403; UTEROGLOBLN_1; FALSE_NEG.
 DR PROSITE: PS00404; UTEROGLOBLN_2; FALSE_NEG.
 KM Steroid-binding; signal.
 FT SIGNAL 1 20
 FT CHAIN 21 112 PROSTATIC STEROID-BINDING PROTEIN C2
 FT MOD_RES 21 21 BLOCKED.
 FT DISULFID 28 28 INTERCHAIN (WITH C3) (PROBABLE).
 FT DISULFID 69 69 INTERCHAIN (WITH C3) (PROBABLE).
 FT DISULFID 92 92 INTERCHAIN (WITH C3) (PROBABLE).
 FT

KW Nuclear protein.
 FT NP_BIND 373 380 ATP (BY SIMILARITY).
 FT SITE 473 476 DEAH BOX.
 FT MUTAGEN 386 386 Y->D: SUPPRESSOR PHENOTYPE.
 FT CONFLICT 698 698 A->R (IN REF. 1).
 SQ SEQUENCE 1071 AA: 121652 MW: 87D4CA418F5DAE1 CRC64;

Query Match 12.1%; Score 68; DB 1; Length 1071;
 Best Local Similarity 25.7%; Pred. No. 37;
 Matches 26; Conservative 19; Mismatches 48; Indels 8; Gaps 3;

QY 7 TDSRLAMKLM-VLMIALLLHCYADSGCKLEDMVEKTIINSIDISPEYKELQ---- 60
 DB 439 TDSCTKLVYTGILRELLDDTDKSCVLAIDENHRSNLTDLGFFKLLARRD 498
 QY 61 -EFIDSDAAEAMGKFKOCFLNOSHRTKFMGMHTYDS 100
 DB 499 LKLLTSATMNAK-KESAFPGNAPFTIIGRTPEVOTIYTS 538

RESULT 9
 TRPC_ACTICA STANDARD; PRT; 268 AA.
 ID TRPC_ACTICA
 AC P00911;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Indole-3-glycerol phosphate synthase (EC 4.1.1.48) (IGPS).
 GN TRPC.
 OS Actinobacter calcoaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 OC Actinobacter.
 ON NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88174326; PubMed=6599977;
 RA Kaplan J.B., Goncharoff P., Seibold A.M., Nichols B.P.;
 RT "Nucleotide sequence of the Actinobacter calcoaceticus trpgdc gene cluster".
 RL Mol. Biol. Evol. 14:456-472(1984).
 CC -1- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1'-deoxy-D-ribose 5-phosphate = 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
 CC -1- PATHWAY: FOURTH STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -1- SIMILARITY: BELONGS TO THE TRPC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M36636; AAA21905.1; -
 CC PIR: A01133; GWKECC.
 CC DR HSSE: P00909; IPII.
 CC DR InterPro: IPR001468; IGPS.
 CC DR Pfam: PF00218; IGPS; 1.
 CC DR Prodom: PD001511; IGPS; 1.
 CC DR PROSITE: PS00614; IGPS; 1.
 CC KW Tryptophan biosynthesis; lyase; Decarboxylase.
 CC SQ SEQUENCE 268 AA: 30216 MW: 8DBAC0505D5A7527 CRC64;

Query Match 12.0%; Score 67.5; DB 1; Length 268;
 Best Local Similarity 26.4%; Pred. No. 9.3;
 Matches 19; Conservative 16; Mismatches 28; Indels 9; Gaps 2;

QY 25 LLHCTADSGCKLEDMVEKTIINSIDISPEYKELQEFIDSDAAEAMGKFKOCFLNOSH 84
 DB 143 LIVACISD--OOLHEMSKTAPEYDL-----DVLVEVHDDDELFRALKLSQCLIGVNN 193

QY 85 RFLKNFGIMMHT 96
 DB 194 RNLKTFVDLMT 205

RESULT 10
 SC10_YEAST STANDARD; PRT; 871 AA.
 ID SC10_YEAST
 AC Q06245; P87329;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Exocyst complex component SEC10.
 GN SEC10 OR YLR166C OR L9362.12.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 ON NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 347-353 AND 856-867.
 RX MEDLINE=97133276; PubMed=8976675;
 RA Terbush D.R., Maurice T., Routh D., Novick P.;
 RT "The Exocyst is a multiprotein complex required for exocytosis in Saccharomyces cerevisiae".
 RL EMBO J. 15:6483-6494(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=528BC / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
 RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Menezes S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
 RA Tach A., Trevasaki E., Vignati D., Wilcox L., Woldman P., Vaudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (Mar-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR EXOCYTOSIS.
 CC -1- SUBUNIT: SEC3, SEC5, SEC6, SEC8, SEC10, SEC15 AND EXO70 ARE COMPONENT OF EXOCYST COMPLEX.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: Y08789; CAA70041.1; -
 CC DR EMBL: U51921; AAB67490.1; -
 CC DR EMBL: U17246; AAB67476.2; -
 CC SCD: S0004156; SEC10.
 CC KW Transport; Protein transport; Golgi stack; Coiled coil.
 CC FT DOMAIN 74 101 COILED COIL (POTENTIAL).
 CC SQ SEQUENCE 871 AA: 100341 MW: A32073099815388D CRC64;

Query Match 12.0%; Score 67.5; DB 1; Length 871;
 Best Local Similarity 28.4%; Pred. No. 33;
 Matches 25; Conservative 17; Mismatches 29; Indels 17; Gaps 4;

QY 4 RLNTSSRLAMKLMVLMIALLLHCYADSGCKLEDMVEKTIINSIDISPEYKELQ-EF 62
 DB 190 KLNKLSVKL-MKNLLTL-----SSKLETSSIPKTIINTLVIEKYSKEMENEL 235

QY 63 IDSDAAEAMGKFKOCFLNOSHRTKFMGMHTYDS 90
 DB 236 LENFNSAYRENFTK-LNEIATILLNNF 263

RESULT 11
 RAD_C_HAEIN STANDARD; PRT; 234 AA.
 ID RAD_C_HAEIN

```

AC P44952;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein radc homolog.
GN
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
NCBI_TaxID=727;
RN
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=93350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kellaway A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McEnaney K., Sutton G., Fitzhugh W., Field C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uitterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Plue L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT *Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.;
RL Science 269:496-512(1995).
CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE RADC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32776; AAC22613.1; -.
CC DR TIGR: H10952; -.
CC DR InterPro: IPR001405; RadC.
CC DR ProDom: PD007415; RadC; 1.
CC DR PROSITE: PS01302; RADC; 1.
CC KM DNA repair; Complete proteome.
CC SQ SEQUENCE 234 AA; 26774 MW; 25472EB289E5DA69 CRC64;

Query Match 11.8%; Score 66.5; DB 1; Length 234;
Best Local Similarity 26.6%; Pred. No. 10;
Matches 21; Conservative 15; Mismatches 26; Indels 17; Gaps 3;

OY 37 LLEDVVEKTIINSISIPYKELLOEFIDSDAAEAMGKFCFLNOSHRLTKNFGIMHT 96
DB 103 LKQDMSTPIIND--PEYKVL--FLTELOHEEREVWVFLDNQHLIKERFLDT 156
OY 97 VY-----DSIWCN 104
DB 157 IYSAVYPREIIEKALYCN 175

RESULT 12
TAZI_SCHPO STANDARD; PRT; 663 AA.
ID TAZI_SCHPO
AC P79005; 000049;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Telomere length regulator tazi.
GN TAZI OR MYB1 OR MYB OR SPAC16A10.07C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;

```

```

RN
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=97186578; PubMed=9034194;
RA Cooper J.P., Nimmo E.R., Allshire R.C., Cech T.R.;
RT *Regulation of telomere length and function by a Myb-domain protein
RT in fission yeast.;
RL Nature 385:744-747(1997).
RN
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Wood V., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases
CC -1- FUNCTION: REGULATES TELOMERE LENGTH AND FUNCTION. REQUIRED FOR THE
CC REPRESSION OF TELOMERE-ADJACENT GENE EXPRESSION AND FOR NORMAL
CC MEIOSIS OR SPOULATION. IT MAY BE A NEGATIVE REGULATOR OF THE
CC TELOMERE-REPLICATING ENZYME, TELOMERASE, OR MAY PROTECT AGAINST
CC ACTIVATION OF TELOMERASE-INDEPENDENT PATHWAYS OF TELOMERE
CC ELONGATION. IT MAY BE INVOLVED IN THE INTERACTIONS BETWEEN
CC CHROMOSOMES AND SPINDLE PROTEINS, DISRUPTION OF THESE
CC INTERACTIONS WOULD LEAD TO DEFECTIVE MEIOSIS.
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09406; CA470568.1; -.
CC DR EMBL: 297185; CAB10000.1; -.
CC DR InterPro: IPR001005; Myb-DNA_bind.
CC DR Pfam: Pf00249; myb-DNA-binding; 1.
CC DR SMART: SM00395; SANT; 1.
CC DR PROSITE: PS00037; MYB_1; 1.
CC DR PROSITE: PS00334; MYB_2; FALSE NEG.
CC DR PROSITE: PS50090; MYB_3; FALSE NEG.
CC KM Telomere; Nuclear protein; DNA-binding; Repeat.
CC FT DNA_BIND 556 612 MYB MYB
CC SQ SEQUENCE 663 AA; 74647 MW; 5852P21518031152 CRC64;

Query Match 11.8%; Score 66.5; DB 1; Length 663;
Best Local Similarity 21.0%; Pred. No. 31;
Matches 22; Conservative 28; Mismatches 38; Indels 17; Gaps 5;

OY 8 DSSRLAKLIMVIMIALLLHCYADSGCKLHIDVETKINSISIPYKELLOEFIDSDA 67
DB 119 DONMAIVRSIMDIKASLV-----NDCONTANNAEOKVMVNSAI--FSSESKDIYNPES 171
OY 68 AAEAMGK--FPCFLNOSHRLTKNFGIMHTV-----YDSIWC 103
DB 172 FSRIGKRVKDYVFNQDLITK-VGLEFRTLPFSYIKKYDAYVC 215

RESULT 13
IL15_MOUSE STANDARD; PRT; 162 AA.
ID IL15_MOUSE
AC P48346;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Interleukin-15 precursor (IL-15).
GN IL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN
[1]
RP SEQUENCE FROM N.A.

```

SRRAIN-WC/REI X C57BL/6J; TISSUE-Bone marrow;
 RA MEDLINE-95278940; Pubmed-7759105;
 RA Anderson D.M., Johnson L., Giaccum M.B., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Valentine V., Kirstein M.N., Shapiro D.N., Morris S.W.,
 RA Grubstein K., Cosman D.;
 RT *Chromosomal assignment and genomic structure of IL15.*;
 CC Genomics 25:701-706(1995).
 CC
 CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
 CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
 CC WITH COMPONENTS OF IL-2R, INCLUDING IL-2R BETA AND PROBABLY IL-2R
 CC GAMMA BUT NOT IL-2R ALPHA.
 CC
 CC -1- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U14332; AAA5377.1; -
 CC DR MGD: MGI:103014; IL15.
 CC DR Interpro: IPR003443; Interleukin_15.
 CC DR Pfam: PF02372; IL15; 1.
 CC KW Cytokine; glycoprotein; Signal.
 CC FT SIGNAL 1 29 POTENTIAL.
 CC FT PROPEP 30 48 POTENTIAL.
 CC FT CHAIN 49 162 INTERLEUKIN-15.
 CC FT DISULFID 83 133 POTENTIAL.
 CC FT DISULFID 90 136 POTENTIAL.
 CC FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 108 108 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 162 AA; 18593 MW; 68C971496CBE296 CRC64;
 CC
 CC Query Match 11.6%; Score 65.5; DB 1; Length 162;
 CC Best Local Similarity 34.0%; Pred. No. 8.6;
 CC Matches 18; Conservative 6; Mismatches 10; Indels 19; Gaps 2;
 CC
 CC QY 31 ADSGCKLEDVYEKINSIDISIPYKELLQEFIDSDAAAEAMGKRCOCLNS 83
 CC DB 129 AESGCKCEELFEKFT-----PFEPLQSF-----RIVQMPINFS 162
 CC
 CC RESULT 14
 CC VE1_HPV08 STANDARD: PRT: 603 AA.
 CC ID VE1_HPV08
 CC AC P06420;
 CC DT 01-JAN-1988 (Rel. 06, Created)
 CC DT 01-JAN-1988 (Rel. 06, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE Replication protein E1.
 CC GN E1.
 CC OS Human papillomavirus type 8.
 CC OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 CC OC Papillomavirus.
 CC OX NCBI_TaxID=10579;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE-86200410; Pubmed-3009874;
 CC RA Fuchs P.G., Iftner T., Weininger J., Pfister H.,
 CC RT *Epidermodysplasia verruciformis-associated human papillomavirus 8:
 CC RT genomic sequence and comparative analysis.*;
 CC RL J. Virol. 58:626-634(1986).
 CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
 CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
 CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
 CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M12737; -; NOT_ANNOTATED_CDS.
 CC DR PIR: A03656; M1WLB.
 CC DR Interpro: IPR001177; Papillom_E1.
 CC DR Pfam: PF00519; E1; 1.
 CC DR Pfam: PF00524; E1.N; 1.
 CC KW Early protein: DNA replication; Helicase; ATP-binding; DNA-binding;
 CC Nuclear protein.
 CC FT NUC_BIND 431 438 ATP (POTENTIAL).
 CC FT BIND 431 438
 CC SQ SEQUENCE 603 AA; 68821 MW; 0813860098DA8A8D CRC64;
 CC
 CC Query Match 11.6%; Score 65.5; DB 1; Length 603;
 CC Best Local Similarity 22.2%; Pred. No. 36;
 CC Matches 22; Conservative 21; Mismatches 23; Indels 33; Gaps 4;
 CC
 CC QY 32 DSGCKLEDVYEKINSIDISIP-----EYKELLQEFIDSDAAAEAMGK 74
 CC DB 112 DSGVELLNNEADYSHVEVPAIDSRPEDEGGSGALDIDYALR--SSNTKATLMK 168
 CC
 CC QY 75 FKOCFLN-----QSHRTLNFGIM---MRTYDS 100
 CC DB 169 FKFAFGDFNELLRPFKSKYTCGCNWWVAVAVHDYES 207
 CC
 CC RESULT 15
 CC GLUT1_YEAST STANDARD: PRT: 2144 AA.
 CC ID GLUT1_YEAST
 CC AC Q12680; Q12290;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE Glutamate synthase [NADPH] precursor (EC 1.4.1.13) (NADPH-GOGAT).
 CC GN GLUT1 OR YDL171C.
 CC OS Saccharomyces cerevisiae (Baker's yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CC OX NCBI_TaxID=4932;
 CC RN [1]
 CC RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 CC RX STRAIN-CN36;
 CC RX MEDLINE-97082505; Pubmed-8923741;
 CC RA Fillelci P., Martegani M.P., Valenzuela L., Gonzalez A., Ballario P.,
 CC RT *Sequence of the GLUT1 gene from Saccharomyces cerevisiae reveals the
 CC RT domain structure of yeast glutamate synthase.*;
 CC RL Yeast 12:1359-1366(1996).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN-S288C;
 CC RA Pohl T.M.;
 CC RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NADP(+) = L-glutamine + 2-
 CC OXOGlutamate + NADPH.
 CC -1- COFACTOR: BINDS A 3FE-4S CLUSTER: FAD AND FMN.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X89221; CAA61505.1; -

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:02:09 ; Search time 25.21 Seconds
(without alignments)
741.113 Million cell updates/sec

Title: US-09-673-395A-238
Perfect score: 564
Sequence: 1 CHARLNTDSSRLAKMLMLV.....NFGIMHTVYDSIWCNKNKS 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriap:*
17: sp._archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	283	50.2	93	6 09GK63	09GK63 oryctolagus
2	232	41.1	93	6 09GK61	09GK61 oryctolagus
3	229	40.6	93	6 09GK64	09GK64 oryctolagus
4	227	40.2	93	6 09GK62	09GK62 oryctolagus
5	191.5	34.0	95	11 09JHB9	09JHB9 ratusus norv
6	178.5	30.2	94	11 09QXF3	09QXF3 mesocricetu
7	170.5	30.2	92	11 09QXF2	09QXF2 mesocricetu
8	77.5	13.7	92	11 09IWB5	09IWB5 mus musculu
9	76.5	13.6	280	3 013997	013997 schizosacch
10	76	13.5	212	16 09A622	09A622 caulobacter
11	73.5	13.0	590	10 09CAN6	09CAN6 arabidopsis
12	73	12.9	702	5 09VH96	09VH96 drosophila
13	73	12.9	1688	13 057483	057483 rana catesb
14	71.5	12.7	487	10 09LY86	09LY86 arabidopsis
15	71.5	12.7	4848	2 007944	007944 streptomyce
16	71	12.6	303	16 09ZGB8	09ZGB8 rickettsia

17	71	12.6	530	5 09XZ83	09XZ83 amoeba prot
18	71	12.6	828	5 09VC06	09VC06 drosophila
19	71	12.6	1187	5 09V7N3	09V7N3 drosophila
20	71	12.6	1189	5 09SR60	09SR60 drosophila
21	70.5	12.5	343	5 09W212	09W212 drosophila
22	70	12.4	457	5 018534	018534 caenorhabd
23	69.5	12.3	108	9 09JML7	09JML7 bacterioph
24	69.5	12.3	412	10 09M2A1	09M2A1 arabidopsis
25	69.5	12.3	621	10 09ASZ8	09ASZ8 arabidopsis
26	69.5	12.3	735	10 09LN85	09LN85 arabidopsis
27	69.5	12.3	1048	10 09SCV7	09SCV7 arabidopsis
28	69	12.2	219	17 028785	028785 archaeoglob
29	69	12.2	638	10 09M393	09M393 arabidopsis
30	69	12.2	799	2 09L826	09L826 enterococu
31	68.5	12.1	173	3 081626	081626 arabidopsis
32	68.5	12.1	213	16 097M18	097M18 clostridium
33	68.5	12.1	384	5 09VG66	09VG66 drosophila
34	68.5	12.1	496	4 09NMG0	09NMG0 homo sapien
35	68.5	12.1	619	4 09HDC6	09HDC6 homo sapien
36	68.5	12.1	694	4 09UPP0	09UPP0 homo sapien
37	68.5	12.1	720	4 09NR80	09NR80 homo sapien
38	68	12.1	516	16 09A9K8	09A9K8 caulobacter
39	68	12.1	570	12 09DHR3	09DHR3 yaba-like d
40	68	12.1	4717	3 094248	094248 schizosacch
41	67	11.9	738	10 082380	082380 arabidopsis
42	67	11.9	1001	15 09WC54	09WC54 human immun
43	66.5	11.8	126	5 027011	027011 tenebrio mo
44	66.5	11.8	199	13 090ZB3	090ZB3 xenopus lae
45	66.5	11.8	274	16 09PFN6	09PFN6 xyella las

ALIGNMENTS

RESULT 1
ID 09GK63 PRELIMINARY; PRT; 93 AA.
AC 09GK63:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LIPOPHILIN CL2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LACRIMAL GLAND;
RA Zhao C., Nguyen T. X., Lehrer R. I.;
RT "Rabbit Lipophilins.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF308618; AMG42806.1; -;
DR InterPro: IPR003627; Mammaglobn_prostn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln_1.
DR ProDom: PD029354; Mammaglobn_prostn; 1.
DR PROSITE: PS000354; Mammaglobn.
SQ SEQUENCE 93 AA; 10456 MW; 4915E45D9134AE91 CRC64;

Query Match 50.2%; Score 283; DB 6; Length 93;
Best Local Similarity 57.8%; Pred. No. 1, le-22;
Matches 52; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY	14	MLKLMLLALALILHCYADSGCKLLEDVETKINSISIPYKELQEFIDSNAAAEAG 73
DB	1	MKVVMVLLALALPLCYAGSGCVLLESVVKTTIDPSVSEYKADLORFIDTQTBAAVF 60
QY	74	KFKCFLNOSHRTAKNFGIMHTVYDSIWC 103
DB	61	EFKCFLSQSNSETLANFRVMHTYDSLYC 90


```

RESULT      2
09GK61
ID 09GK61      PRELIMINARY:      PRT:      93 AA.
AC 09GK61;
DT 01-MAR-2001 (TREMBLrel, 16, Created)
DT 01-MAR-2001 (TREMBLrel, 16, last sequence update)
DT 01-JUN-2001 (TREMBLrel, 17, last annotation update)
DE LIPOPHILIN CS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN      [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SUBMAXILLARY;
RA Zhao C., Nguyen T.X., Lehrer R.I.:
RT "Rabbit Lipophilins ".
RL Submitted (SEP-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF308620; AAC42808.1; -.
DR InterPro; IPR000362; Mammaglobn_prostin.
DR InterPro; IPR000329; Uteroglobln.
DR Pfam; PF01099; uteroglobln; 1.
DR PRINTS; P000486; UTEROGLOBIN.
DR PRODOM; PD028354; Mammaglobn_prostin; 1.
DR SMART; SM00096; UTR; 1.
SQ SEQUENCE 93 AA; 10509 MW; E2B015AD319B2249 CRC64;

```

Query Match	41.18;	Score 232;	DB 6;	Length 93;
Best Local Similarity	50.08;	Pred. No. 3e-17;	25;	
Matches 45;	Conservative			0;
				Gaps 0;

QY 14 MLLMVLMLAALLHCAAGSGCKLLLEDMEKTLINSDISIPYKELDIFSDSAAAEEAG 73
 QY 1 MLLVLMVLMALPLPYCAAGSCFPFEKAKMTLLNSNVSTAEIIVLYKNIYNDRETELAV 60
 Db 74 KFKQCFLNQSHRTLKNPGLMHTTVYDSIAC 103
 QY 61 EFKNCFISQSEETLRNVKEMETIYNSKIC 90
 Db

```

RESULT 3
Q9GK64 PRELIMINARY; PRT; 93 AA.
Q9GK64 AC Q9GK64:
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE LIPOPHILIN CL.
OC Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxId=9986;
RN 11]
RP SEQUENCE FROM N.A.
RC TISSUE=LACRIMAL GLAND;
RA Zhao C., Nguyen T.X., Lehrer R.I.;
RT "Rabbit Lipophilins.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
EMBL AF308611; AAG42805.1; -.
DR InterPro: IPR003627; Mammaglobin_protstn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.1.
DR ProDom: PD029354; Mammaglobn_protstn.1.
SQ SEQUENCE 93 AA: 10350 MW; 9F9206C4437280AD CRC64;

```

Query Match	40.6%	Score 229	DB 6	Length 93	
Best Local Similarity	45.6%	Pred. No. 6	2e-17		
Matches 41	Conservative 26	Mismatches 23	Indels 0	Gaps 0	
OY	14	MKLLMTVLAALLHCHTADSCCKLLEMDVKEKTIINSDSIPREYKELQEPIDSDAAAEAWG	73		
	: :	: : : :	: :	: :	: :

Db 1 MKMVLVLATLPFYCAGSGCVIIIESVLDKTDIPSVSEDTYYLQKRIILTDAAVALE 60

Qy 74 KFKCETLNQSHRTLKNGFLMHHYVDISMC 103
: |||||::|| : : : ||::||
61 ELKQCFLSQSNETLANKYVEYAVFSLYC 90

```

RESULT 4
09GK62 PRELIMINARY: PRT: 93 AA.
AC 09GK62:
DT 01-MAR-2001 (TEMBIrel. 16, Created)
DT 01-MAR-2001 (TEMBIrel. 16, last sequence update)
DT 01-JUN-2001 (TEMBIrel. 17, last annotation update)
DE LIPOPHILIN CP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 11]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Zhao C., Nguyen T.X., Lehrer R.I.;
RT "Rabbit Lipophilins.";
RL Submitted (SEP-2000) to the EMBI/GenBank/DBJ databases.
DR EMBL; AF308619; AAC42807.1; -.
DR InterPro: IPR003627; Mammagln_protn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; uteroglobln.1.
DR Prodom: PD029354; Mammagln_protn.1.
SQ SEQUENCE 93 AA; 10332 MW; C9DC35B17D372F32 CRC64;

```

Query Match	40.28;	Score 227;	DB 6;	Length 93;
Best Local Similarity	50.08;	Pred. No. 1e-16;		
Matches 45; Conservative	16;	Mismatches 29;	Indels 0;	Caps 0;

[illegible]

RESULT	5		
ID	Q9JHH9	PRELIMINARY:	PRT: 95 AA.
AC	Q9JHH9;		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)		
DT	01-OCT-2001 (TREMBLrel. 18, last annotation update)		
DE	PROSTATIC SPHEROID BINDING PROTEIN.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCBI_Taxid-10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-84057754; PubMed-6665625;		
RA	Hurst H.C., Parter M.G.;		
RT	"Rat prostatic steroid binding protein: DNA sequence and transcript		
RL	maps of the two C3 genes.";		
EMBL	EMBO J. 2:769-774(1983).		
DR	EMBL; V01260; CAB75892.1;		
DR	EMBL; V01261; CAB75892.1; JOINED.		
DR	EMBL; V01262; CAB75892.1; JOINED.		
DR	InterPro: IPR003627; Mammaglobn.prostln.		
DR	InterPro: IPR000329; Uteroglobln.		
DR	Pfam: PF01099; Uteroglobln. 1.		
DR	ProDom: PD029354; Mammaglobn.Prostln. 1		
DR	SK09ENE 95 AA; 10622 MW; B209F1BE177C5244 CRC64;		

```

OX NCHI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LAKE VIEW; TISSUE-HARDERIAN GLAND;
RA Alvarez J., Alder J., Ashman K., Martin Alonso J M., Dominguez P.;
RT "Cloning and differential expression of two isoforms of hecroglobin,"
RL a novel heterodimeric glycoprotein of the uteroglobin gene family.";
DR Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ252139; CAB64661.1;
DR InterPro: IPR003627; Mammaglobin_protln.
DR InterPro: IPR000329; uteroglobin.
DR Pfam: Pf01099; uteroglobin_1.
DR Pfam: PD029354; Mammaglobin_protctn; 1.
KW Signal.
FT SIGNAL 1 16 POTENTIAL.
FT CHAIN 17 94 HETEROGLOBIN B2 SUBUNIT.
FT SEQUENCE 94 AA; 10821 MW; 16C640C0674224C9 CRC64;
SQ

Query Match 30.2%; Score 170.5; DB 11; Length 94;
Best Local Similarity 37.0%; Pred. No. 1e-10; Gaps 1;
Matches 34; Conservative 23; Mismatches 34; Indels 1; Gaps 1;

OY 14 MKLVLVLMIALALHCYAD-SGCKLEIDPWVEKTTISDSISPEYKELLOEFTDSAAAKAM 72
DB 1 MLVIVFMTALIPVCRINSSCCNMDDAIATINSSVMEERYHETVKKYTTLPYIRSTV 60
OY 73 GKFKOCFLNOSHRTLKNGLMNHTYYDSIWCN 104
DB 61 EKFEKCFAKOSMDTOHNIFFVWYAVYNSDKS 92

RESULT 8
OY1WB5 PRELIMINARY: PRT: 92 AA.
ID OY1WB5
AC OY1WB5;
DT 01-DEC-2001 (Tremblrel, 19, Created)
DT 01-DEC-2001 (Tremblrel, 19, last sequence update)
DT 01-DEC-2001 (Tremblrel, 19, last annotation update)
DE HYPOTHETICAL 10.1 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCHI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SALIVARY GLAND;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
KW EMBL: BC016132; AAB16132.1; -.
KW Hypothetical protein.
SQ SEQUENCE 92 AA; 10080 MW; F1A7557E0F5568BB CRC64;

Query Match 13.7%; Score 77.5; DB 11; Length 92;
Best Local Similarity 36.4%; Pred. No 0.75;
Matches 28; Conservative 10; Mismatches 30; Indels 9; Gaps 4;

OY 14 MKLMLVLM--ALLLHCYADSG-CKLLEDVEXKTINSISIPYKELLOEFTDSAAAE 70
DB 1 MKLGLALLLGAALLITSEGDGCLPALQRRVDILFNG--TTEEYVAYLKFQENMKVILK 58
OY 71 AMGFKOCFLNOSHRTL 87
DB 59 NAANIKKC---SDRTL 71

RESULT 9
OY13997 PRELIMINARY: PRT: 280 AA.
ID OY13997
AC OY13997;
DT 01-JUN-1998 (Tremblrel, 06, Created)
DT 01-JUN-1998 (Tremblrel, 06, last sequence update)

```


01-MAY-2000 (TrEMBLrel. 13, last sequence update)
 01-DEC-2001 (TrEMBLrel. 19, last annotation update)
 CG8358 PROTEIN.
 GN CG8358.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC EBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaratunga P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer H.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt C., Andrews-Pfannkoch C., Baldwin D.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beson K.Y., Bess P.V., Bernier B.P., Bhandari D., Bolintinas S.,
 RA Borokov D., Botchan M.R., Bouck J., Brodeur G., Center A., Chandra I.,
 RA Butts K.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel K., Gabrielian A.E., Garb N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegam C.,
 RA Jallat B.E., Kalush C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laso P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mettel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reliford K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Stong R., Sun E.,
 RA Svikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Weissman D.A., Weisslock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003683; AAF54423.1;
 DR MEROPS: M13.0PM;
 DR F1ybase: Fpnu0037727; CG8358.
 DR InterPro: IPR000718; Peptidase_M13.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01431; Peptidase_M13; 1.
 DR PRINTS: PR00766; NEPRILYSIN.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN.1.
 DR SEQUENCE 702 AA; 80825 MW; 4FCA1A39F1F0BB6 CRC64;

Query Match 12.9%; Score 73; DB 5; Length 702;
 Best Local Similarity 27.8%; Pred. No. 23;
 Matches 25; Conservative 19; Mismatches 20; Indels 26; Gaps 4;

0Y 8 DSSRLAKLMLVLAALILHYCAADSCKLEMEVETINSDI-----SIPEYELLQ 60
 DB 4 ESILKLSIMLWIVAAVLTDCARS-----LVDRSENSDNGSSTNSAAEYO--LQ 54
 0Y 61 EFIDSDAAAEAMGKFCOCFETNOSHTLKNF 90
 DB 55 EY-----AEFMKSYNOSVCEPCNF 74

RESULT 13
 ID 057483 PRELIMINARY: PRT: 1688 AA.
 AC 057483;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, last annotation update)
 DE VOLTAGE-DEPENDENT L-TYPE CALCIUM CHANNEL, ALPHA-1S SUBUNIT
 DE (FGALPHA1S).
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
 NC EBI_Taxid=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SKELTAL MUSCLE;
 RX MEDLINE=98411359; PubMed=9738021;
 RA Zhou J., Cribbs L., Yi J., Shirokov R., Perez-Reyes E., Rios E.;
 RT "Molecular cloning and functional expression of a skeletal muscle
 dihydropyridine receptor from Rana catesbeiana.";
 RL J. Biol. Chem. 273:25503-25509(1998).
 CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDATE THE
 ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
 IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
 CONTRACTION, GENE EXPRESSION, CELL MOTILITY, CELL DIVISION AND
 CELL DEATH. THE ISOFORM ALPHA-1S GIVES RISE TO L-TYPE CALCIUM
 CURRENTS.
 CC -1- FUNCTION: LONG-LASTING (L-TYPE) CALCIUM CHANNELS BELONG TO THE
 "HIGH-VOLTAGE ACTIVATED (HVA)" GROUP. THEY ARE BLOCKED BY
 DIHYDROPYRIDINES (DHP), PHENYLAALKYLAMINES, BENZOTHAZEPINES, AND B
 OMEGA-AGATOXIN-IIIA (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE
 TO OMEGA-COROTOXIN-GVIA (OMEGA-CITX-GVIA) AND OMEGA-AGATOXIN-IVA
 (OMEGA-AGA-IVA). CALCIUM CHANNELS CONTAINING THE ALPHA-1S SUBUNIT
 PLAY AN IMPORTANT ROLE IN EXCITATION-CONTRACTION COUPLING IN
 SKELETAL MUSCLE (BY SIMILARITY).
 CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE FOR-
 FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY. AN
 ADDITIONAL GAMMA SUBUNIT IS PRESENT ONLY IN SKELETAL MUSCLE L-TYPE
 CHANNEL (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE SPECIFIC.
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -1- DOMAIN: THE LOOP BETWEEN REPEATS II AND III INTERACTS WITH THE
 RYANODINE RECEPTOR, AND IS THEREFORE IMPORTANT FOR CALCIUM RELEASE
 FROM THE ENDOPLASMIC RETICULUM NECESSARY FOR MUSCLE CONTRACTION
 (BY SIMILARITY).
 CC -1- PTM: PHOSPHORYLATION BY CAPK STIMULATES THE CALCIUM CHANNEL
 FUNCTION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
 FAMILY.
 DR EMBL: AF037625; AAC36126.1;
 DR InterPro: IPR000636; Cation_channel_nou_119.
 DR InterPro: IPR002111; Cat_channel_TripL.
 DR InterPro: IPR002077; Ca_channel.
 DR InterPro: IPR001682; Channel_pore_Ca_Na.
 DR InterPro: IPR001064; Crystal11in.
 DR Pfam: PF00520; ion_trans_4.
 DR PRINTS: PR00167; CACCHANNEL.
 DR PROSITE: PS00225; CRISTALLIN_BETAGAMMA; UNKNOWN.1.
 DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium binding; Phosphorylation.

```

FT CARBOHYD 80 80 N-LINKED (GLCNAC, ...) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC, ...) (POTENTIAL).
FT SEQ 1688 AA; 192421 MW; 58BA06B3882017A CRC64;
SQ
Query Match 12.9%; Score 73; DB 13; Length 1688;
Best Local Similarity 24.1%; Pred. No. 64;
Matches 28; Conservative 16; Mismatches 36; Indels 36; Gaps 5;
OY 16 LHMVIMLAALHCHVADSGCKLDEDMVKTI---NSDISIPEYKELLQGFIDSD--AAAE 70
DB 194 LHMIALVLEMTIITAYIGLELFSCKMKRTKTYEDDITAT-----VNEKPACRS 244
OY 71 AMGFKQCEFLNOSH-----RTKNFSLMHTVYDSI-----WCN 104
DB 245 STGGHQCSTINSGECGMMWPGPNNGITFHDNFGPAMLTVCITMGMTVEVLYWVN 300
RESULT 14
ID 09LV86 PRELIMINARY; PRT; 487 AA.
AC 09LV86;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, P1 CLONE:MXR3.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotsids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Kaloh T., Asamini E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RL clones."
RL DNA REFL. 7:31-63(2000).
DR EMBL: AB019236; BAA97303.1; -.
SQ SEQUENCE 487 AA; 54862 MW; 632BAVE298807D27 CRC64;
Query Match 12.7%; Score 71.5; DB 10; Length 487;
Best Local Similarity 22.7%; Pred. No. 22;
Matches 17; Conservative 16; Mismatches 33; Indels 9; Gaps 1.
OY 27 LHCYADSGCKLDEDMVKTI NSDISIPEYKELLQGFIDS-----DAAAEAMGKFKQ 77
DB 366 VCECEYDMWKSVELEFDTSLINEVLLPQKAVAFEEFKESDKAMAAVDTAQAELERAKE 425
OY 78 CFLNOSHRTLNKFG 92
DB 426 GLSEETKRAYQEMRL 440
RESULT 15
ID 007944 PRELIMINARY; PRT; 4848 AA.
AC 007944;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE PRISTINAMYCIN I SYNTHASE 3 AND 4.
GN SNBDE.
OS Streptomyces pristinaespiralis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=38300;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-SP92;
 RA de Crecy-Lagard V.A., Saurin W., Thibaut D., Gil P., Naudin L.,
 RA Crouzet J., Blanc V.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP92;
 RX MEDLINE-97158664; Pubmed-9006024;
 RA De Crecy-Lagard V., Blanc V., Gil P., Naudin L., Lorenzon S.,
 RA Farnecion A., Bamas-Jacques N., Crouzet J., Thibaut D.;
 RT *Pristinamycin I biosynthesis in Streptomyces pristinaespiralis:
 RT molecular characterization of the first two structural peptide
 RT synthetase genes.*;
 RL J. Bacteriol. 179:705-713(1997).
 DR EMBL: Y11548; CAA72312.1; -;
 DR EMBL: X98690; CAA67249.1; -;
 DR HSSP; P14687; 1AMU
 DR InterPro: IPR002106; AA_LNA_Ligase_II.
 DR InterPro: IPR000873; AMP-bind.
 DR Interf 3; IPR01242; DUF4.
 DR InterPro: IPR002336; Erythecurin.
 DR InterPro: IPR000379; Est_1ip_thioest_actsite.
 DR InterPro: IPR001601; Meth-transf.
 DR InterPro: IPR003880; Phosphopant_attach.
 DR InterPro: IPR000051; SAM_bind.
 DR InterPro: IPR01031; Thioesterase.
 DR Pfam: PF00501; AMP-binding; 4.
 DR Pfam: PF00668; Condensation; 5.
 DR Pfam: PF00550; PP-binding; 4.
 DR Pfam: PF00975; Thioesterase; 1.
 DR PRINTS: PR00154; AMPBINDING.
 DR PRINTS: PR00611; ERYTHECURIN.
 DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 DR PROSITE: PS50075; ACP_DOMAIN; 4.
 DR PROSITE: PS00455; AMP_BINDING; 4.
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_4.
 KM Phosphopantetheine.
 SQ SEQUENCE 4848 AA; 522081 MW; E17591617A2B9A0E CRC64;

Query Match 12.7%; Score 71.5; DB 2; Length 4848;
 Best Local Similarity 32.6%; Pred. No. 3.1e+02;
 Matches 29; Conservative 10; Mismatches 31; Indels 19; Gaps 3;
 QY 2 HARLMTDSSRLAMKLLMYLM--LAALLHGYADSCCKL-----LEDMVEKPTINS 48
 DB 2716 HORLMAALTEGRASLFMVQAAFAALLTRHCAGTIDPLGSPACRTDADLEDVNGFVYNT 2775
 QY 49 DI-----SIPYKELIOEFIDSDAAAEA 71
 DB 2776 LVKRTDTSGDPTFFRELLQVRRESDLAAYN 2804

Search completed: June 20, 2002, 11:05:59
 Job time: 230 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:01:45 ; Search time 17.99 Seconds
(without alignments)
526.499 Million cell updates/sec

Title: US-09-673-395A-238
Perfect score: 564
Sequence: 1 CHARLNTDSSRLMKLMLV.....NFGLMHTVDSIMCNKSN 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 265615 seqs, 87701152 residues

Total number of hits satisfying chosen parameters: 265615

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending-Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PCF_NEW.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US09_NEW.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US10_NEW.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US50_NEW.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US50_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	87.9	95	1	Sequence 187, App
2	496	87.9	95	6	Sequence 187, App
3	283	50.2	93	6	Sequence 58, Appl
4	291	49.8	93	1	Sequence 503, App
5	281	49.8	93	6	Sequence 493, App
6	281	49.8	410	6	Sequence 494, App
7	281	49.8	410	6	Sequence 494, App
8	281	49.8	743	6	Sequence 493, App
9	281	49.8	743	6	Sequence 493, App
10	281	49.8	1095	6	Sequence 493, App
11	281	49.8	1095	6	Sequence 493, App
12	206	36.5	220	5	Sequence 498, App
13	77	13.7	21	1	Sequence 498, App
14	77	13.7	21	6	Sequence 498, App
15	70	12.4	457	7	Sequence 6769, App
16	68	12.1	1071	7	Sequence 22283, A
17	67.5	12.0	871	7	Sequence 22330, A
18	66	11.7	20	6	Sequence 499, App
19	66	11.7	20	6	Sequence 499, App
20	66	11.7	135	5	Sequence 8278, Ap
21	66	11.7	293	5	Sequence 42, Appl
22	65.5	11.6	429	5	Sequence 1310, Ap
23	64.5	11.4	395	5	Sequence 7865, Ap
24	64.5	11.4	446	5	Sequence 7865, Ap
25	64.5	11.4	610	5	Sequence 7864, Ap
26	64	11.3	1280	5	Sequence 29952, A

27	64	11.3	1310	5	US-09-935-625-29951	Sequence 29951, A
28	64	11.3	1389	5	US-09-935-625-29950	Sequence 29950, A
29	63.5	11.3	246	6	US-10-104-047-3660	Sequence 3660, Ap
30	63.5	11.3	507	6	US-09-540-209B-8498	Sequence 8498, Ap
31	63.5	11.3	1360	5	US-09-573-655B-282	Sequence 282, App
32	63	11.2	410	7	US-60-360-039-10896	Sequence 10896, A
33	62.5	11.1	1094	5	US-09-573-655B-1388	Sequence 1388, Ap
34	62	11.0	193	5	US-09-540-209B-10334	Sequence 10334, A
35	62	11.0	193	5	US-09-540-209B-10352	Sequence 10352, A
36	62	11.0	472	5	US-09-540-209B-5684	Sequence 5684, Ap
37	62	11.0	569	5	US-09-573-655B-216	Sequence 216, App
38	62	11.0	1204	5	US-09-935-625-30799	Sequence 30799, A
39	62	11.0	1219	5	US-09-935-625-30802	Sequence 30802, A
40	62	11.0	1234	5	US-09-935-625-30798	Sequence 30798, A
41	62	11.0	1313	5	US-09-935-625-30797	Sequence 30797, A
42	62	11.0	1384	5	US-09-935-625-30801	Sequence 30801, A
43	61.5	10.9	422	5	US-09-573-655B-1311	Sequence 1311, Ap
44	61.5	10.9	734	7	US-60-360-039-3017	Sequence 3017, Ap
45	61.5	10.9	915	1	PCF-US02-13142-3197	Sequence 3197, Ap

ALIGNMENTS

RESULT 1
PCF-US02-07826-187
Sequence 187, Application PC/TUS0207826
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification, FILE REFERENCE: MRI-0303C
CURRENT APPLICATION NUMBER: PCF/US02/07826
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 187
LENGTH: 95
TYPE: PCT
ORGANISM: Homo sapiens
PCF-US02-07826-187

Query Match 87.9%; Score 496; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 7.7e-47;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 14 MLLVLMALALLHCYADSGCKLLEDAVEKTTNSDISPEKKELLOFIDDAAEAMG 73
|||||
DB 1 MLLVLMALALLHCYADSGCKLLEDAVEKTTNSDISPEKKELLOFIDDAAEAMG 60
|||||
DB 74 KFKOCFLNOSHRTLNKFGMLMHTVDSIMCNKSN 108
|||||
DB 61 KFKOCFLNOSHRTLNKFGMLMHTVDSIMCNKSN 95
|||||

RESULT 2
US-10-097-340-187
Sequence 187, Application US/10097340

```

; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, JR.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATF
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for the Identification,
; FILE REFERENCE: MRI-030
; FILE REFERENCE: Assessment, Prevention, and Therapy of Ovarian Cancer
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-097-340-187

Query Match      87.9%; Score 496; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 7.7e-47;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MKLWVLMALLLHCYADSGCKLLEDVVEKTIINSIDISPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKLWVLMALLLHCYADSGCKLLEDVVEKTIINSIDISPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOSHRTLKNFGMLMHTVYDSIWCNKN 108
DB 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIWCNKN 95

RESULT 3
US-10-119-480-58
; Sequence 58, Application US/10119480
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C1
; CURRENT APPLICATION NUMBER: US/10/119,480
; CURRENT FILING DATE: 2002-04-09
; NUMBER OF SEQ ID NOS: 246
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 58
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-119-480-58

Query Match      50.2%; Score 283; DB 6; Length 93;
Best Local Similarity 57.8%; Pred. No. 1.1e-23;
Matches 52; Conservative 20; Mismatches 18; Indels 0; Gaps 0;

QY 14 MKLWVLMALLLHCYADSGCKLLEDVVEKTIINSIDISPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKVWVLLALALPLKYCYAGSCVLLSEVVEKTIIDPSVSEYKADLOEFIDTEATNAID 60

QY 74 KFKOCFLNOSHRTLKNFGMLMHTVYDSIWC 103
DB 61 EFKECFLSOSNETLANFRVWHTIYDSLXC 90

RESULT 4
PCT-US02-12378-503
; Sequence 503, Application PC/TUS0212378
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Persing, David H.
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Panger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margareta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.47003PC
; CURRENT APPLICATION NUMBER: PCT/US02/12378
; CURRENT FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 503
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US02-12378-503

Query Match      49.8%; Score 281; DB 1; Length 93;
Best Local Similarity 58.7%; Pred. No. 1.8e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 14 MKLWVLMALLLHCYADSGCKLLEDVVEKTIINSIDISPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKLWVLMALLLHCYADSGCKLLEDVVEKTIINSIDISPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOSHRTLKNFGMLMHTVYDSIWCN 105
DB 61 ELKECFLNOSHRTLKNFGMLMHTVYDSIWCN 92

RESULT 5
US-10-124-805-503
```



```

Query Match Similarity      49.8% ; Score 281 ; DB 6 ; Length 1095;
Best Local Similarity      58.7% ; Pred. No. 3.4e+22;
Matches      54 ; Conservative      13 ; Mismatches      25 ; Indels      0 ; Caps      0

QY      14 MKLWVLMALALLHCYADSCCKLLEMYEKRTINSDISIEPKELLQFTIDSDAAAFAMG 73
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      1 MKLWVLMALALASQCHYAGSGCPLELVISKRTINPOVSKTEYKELLQFTIDNATTNAID 60

QY      74 KFKOCFLNOSHRTLLKNGFLMHTVYDSIWCNM 105
          : : | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      61 ELKECFLNQTDFTLLSNVEVFMQLIYDSSIDL 92

RESULT 12
US-09-684-215A-10
; Sequence 10, Application US/09684215A
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation

```

TITLE OF INVENTION: Methods of Using a Mycobacterium tuberculosis Coding
TITLE OF INVENTION: Sequence to Facilitate Stable and High Yield Expression
FILE REFERENCE: 014058-008010US
CURRENT FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: US 60/158,585
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 10
LENGTH: 220
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURES:
OTHER INFORMATION: Description of Artificial Sequence:ra12-human manmaglobin
US-09-684-215/-10

Query Match 36.5%; Score 206; DB 5; Length 220;
Best Local Similarity 52.1%; Pred. No. 7e-15;
Matches 38; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

OY 33 SGCKLEDMYKRTINSIPETKELQETIDSDAAEAMGKROCKCLNOSHRTLNFG 92
Db 147 SGCPLENNVSKRTIPQVSKTEYKELQETIDSDAAEAMGKROCKCLNOSHRTLNFG 206

OY 93 MMHTVYDSIMCNM 105
Db 207 FMOLYDSLCLD 219

RESULT 13
PCT-US02-12378-498
Sequence 498, Application PC/TUS0212378
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margerita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.47003PC
CURRENT APPLICATION NUMBER: PCT/US02/12378
CURRENT FILING DATE: 2002-04-15
NUMBER OF SEQ ID NOS: 627
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 498
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US02-12378-498

Query Match 13.7%; Score 77; DB 1; Length 21;
Best Local Similarity 68.4%; Pred. No. 0.045;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 28 HCYADSGCKLHDMYKRTI 46
Db 3 HCYAGSGCPLENNVSKRTI 21

RESULT 14
US-10-124-805-498
Sequence 498, Application US/10124805
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: Persing, David H.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470032
CURRENT APPLICATION NUMBER: US/10/124,805
CURRENT FILING DATE: 2002-04-15
NUMBER OF SEQ ID NOS: 627
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 498
LENGTH: 21
TYPE: PRT
ORGANISM: Homo sapiens
US-10-124-805-498

Query Match 13.7%; Score 77; DB 6; Length 21;
Best Local Similarity 68.4%; Pred. No. 0.045;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 28 HCYADSGCKLHDMYKRTI 46
Db 3 HCYAGSGCPLENNVSKRTI 21

RESULT 15
US-60-360-039-6769
Sequence 6769, Application US/60360039
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Chen, Xianfeng
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)A
CURRENT APPLICATION NUMBER: US/60/360,039
CURRENT FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 6769
LENGTH: 457
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-60-360-039-6769

Query Match 12.4%; Score 70; DB 7; Length 457;
Best Local Similarity 28.3%; Pred. No. 10;
Matches 30; Conservative 16; Mismatches 32; Indels 28; Gaps 6;

OY 1 CHARLNTDSSPLAMKILMVLMLA-----ALLHGYADSGCKL-----LEDMYKRT-- 45
Db 178 CTEKMEKSKSRRA--YIMYIMLAOFVFPVMAFCYANITYSVLSKRAQOTIRKMYVERTSA 235
OY 46 INSDISIP-----EYKELQETIDSDAAEAMGKROCKCLNOSHRT 86
Db 236 LESSCAFPVHCLBOYENELNEFLDKQF-----KEKORVVLQNRRT 275

Search completed: June 20, 2002, 11:03:31
Job time: 106 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:01:44 ; Search time 107.96 Seconds

(Without alignments)
352.111 Million cell updates/sec

Title: US-09-673-395A-238

Perfect score: 564
Sequence: 1 CHARLNTDSSRLAMKLIWLV.....NFGLMHIVYDSIWMKSN 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US086.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	564	100.0	108	20	US-09-673-395A-238
2	496	87.9	95	1	PCT-US01-09339-3
3	496	87.9	95	1	PCT-US96-03857-6
4	496	87.9	95	1	PCT-US99-22753-2
5	496	87.9	95	13	US-08-951-750-1
6	496	87.9	95	15	US-09-110-716-31
7	496	87.9	95	18	US-09-412-231-12

8	496	87.9	95	18	US-09-471-276-1555	Sequence 1555, App
9	496	87.9	95	19	US-09-549-342A-3	Sequence 3, Appl
10	496	87.9	95	23	US-09-985-911-6	Sequence 6, Appl
11	475	84.2	92	14	US-09-057-719-1075	Sequence 1075, Ap
12	408	72.3	77	15	US-09-110-716-13	Sequence 13, Appl
13	393.5	69.8	76	15	US-09-110-716-40	Sequence 40, Appl
14	390	69.1	74	17	US-09-367-009-3	Sequence 3, Appl
15	343	60.8	71	26	US-60-234-446-797	Sequence 797, App
16	318	56.4	62	26	US-60-160-203-3646	Sequence 3646, Ap
17	281	49.8	93	1	PCT-US01-04439-1	Sequence 1, Appl
18	281	49.8	93	1	PCT-US01-04439-34	Sequence 34, Appl
19	281	49.8	93	1	PCT-US98-17991-2	Sequence 2, Appl
20	281	49.8	93	1	PCT-US99-22616-2	Sequence 2, Appl
21	281	49.8	93	10	US-08-697-106-2	Sequence 2, Appl
22	281	49.8	93	11	US-08-722-304-2	Sequence 2, Appl
23	281	49.8	93	11	US-08-747-547-3	Sequence 3, Appl
24	281	49.8	93	11	US-08-747-547-10	Sequence 10, Appl
25	281	49.8	93	13	US-08-912-149-17	Sequence 17, Appl
26	281	49.8	93	13	US-08-951-750-3	Sequence 3, Appl
27	281	49.8	93	15	US-09-162-622-2	Sequence 2, Appl
28	281	49.8	93	16	US-09-215-818-5	Sequence 5, Appl
29	281	49.8	93	18	US-09-467-602A-5	Sequence 5, Appl
30	281	49.8	93	19	US-09-509-015-17	Sequence 2, Appl
31	281	49.8	93	19	US-09-534-292-17	Sequence 17, Appl
32	281	49.8	93	19	US-09-580-376-27	Sequence 27, Appl
33	281	49.8	93	21	US-09-757-417-27	Sequence 27, Appl
34	281	49.8	93	21	US-09-780-842-1	Sequence 1, Appl
35	281	49.8	93	24	US-10-007-805-503	Sequence 503, App
36	281	49.8	93	24	US-10-042-945-27	Sequence 27, Appl
37	281	49.8	93	22	US-09-834-759-503	Sequence 503, App
38	281	49.8	93	23	US-09-905-673-1	Sequence 1, Appl
39	281	49.8	93	23	US-09-905-673-34	Sequence 34, Appl
40	281	49.8	93	23	US-09-934-054-3	Sequence 3, Appl
41	281	49.8	93	23	US-09-934-054-10	Sequence 10, Appl
42	281	49.8	93	24	US-10-007-805-503	Sequence 503, App
43	281	49.8	93	24	US-10-042-945-27	Sequence 27, Appl
44	281	49.8	93	24	US-10-076-622-503	Sequence 503, App
45	281	49.8	93	24	US-10-096-319-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-673-395A-238
; Sequence 238, Application US/09673395A
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMAN, BERNI
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILASKI, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 238
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-673-395A-238

Query Match 100.0%; Score 564; D: 20; Length 108;
Best Local Similarity 100.0%; Pred. No. 5.2; 57;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CHARLNTDSSRLAMKLIWLVLAALLHICVADSGCKLEDMVKETINSISTEYRELLQ 60
DB 1 CHARLNTDSSRLAMKLIWLVLAALLHICVADSGCKLEDMVKETINSISTEYRELLQ 60

OY 61 EFIDSDAAAEAMGKFCOFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 108
|
Db 61 EFIDSDAAAEAMGKFCOFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 108

RESULT 2
PCT-US01-09339-3
Sequence 3, Application PC/TUS0109339
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Colpitts, Tracey L.
APPLICANT: Russell, John C.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
TITLE OF INVENTION: DETECTING DISEASES OF THE REPRODUCTIVE TISSUES
FILE REFERENCE: 5972 US P6
CURRENT APPLICATION NUMBER: PCT/US01/09339
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: US 09/467,602
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: US 09/215,818
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: US 08/912,276
PRIOR FILING DATE: 1997-08-17
PRIOR APPLICATION NUMBER: US 08/697,105
PRIOR FILING DATE: 1996-08-19
PRIOR APPLICATION NUMBER: US 08/912,149
PRIOR FILING DATE: 1997-08-15
PRIOR APPLICATION NUMBER: US 08/697,106
PRIOR FILING DATE: 1996-08-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 95
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-09339-3

Query Match 87.9%: Score 496; DB 1; Length 95;
Best Local Similarity 100.0%: Pred. No. 3.4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 73
|
Db 1 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 60
OY 74 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 108
|
Db 61 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 95

RESULT 3
PCT-US96-03857-6
Sequence 6, Application PC/TUS9603857
GENERAL INFORMATION:
APPLICANT: Geitz, Reiner
TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC
TITLE OF INVENTION: STEROID-BINDING FACTOR I, II AND III
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN, CROCHET,
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068-1739
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03857
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-520
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-03857-6

Query Match 87.9%: Score 496; DB 1; Length 95;
Best Local Similarity 100.0%: Pred. No. 3.4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 14 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 73
|
Db 1 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 60
OY 74 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 108
|
Db 61 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 95

RESULT 4
PCT-US99-22753-2
Sequence 2, Application PC/TUS9922753
GENERAL INFORMATION:
APPLICANT: Macina, Roberto A.
TITLE OF INVENTION: A Novel Method of Diagnosing, Monitoring, Staging,
TITLE OF INVENTION: Imaging and Treating Gynecologic Cancers
FILE REFERENCE: DEX-0045
CURRENT APPLICATION NUMBER: PCT/US99/22753
CURRENT FILING DATE: 1999-09-30
EARLIER APPLICATION NUMBER: 60/102,743
EARLIER FILING DATE: 1998-10-02
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 95
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US99-22753-2

Query Match 87.9%: Score 496; DB 1; Length 95;
Best Local Similarity 100.0%: Pred. No. 3.4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 73
|
Db 1 MKLMLVLMIAALLHCYADSGCKLLEDMVEKTI NSDISIPEYKELLOEFIDSDAAAEAMG 60
OY 74 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 108
|
Db 61 KFKOCFLNOSHRTLNFGMLMHTVYDSIWCNMKSN 95

RESULT 5
US-08-951-750-1
Sequence 1, Application US/08951750
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.

```

; APPLICANT: Shah, Purvi
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: NOVEL MAMMOGLOBIN HOMOLOG
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; City: Palo Alto,
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/951,750
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0381 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTN0105
; CLONE: 2295453
;
US-08-951-750-1

Query Match      87.9%; Score 496; DB 13; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 73
      |||||||
DB 1 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 60

OY 74 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 108
      |||||||
DB 61 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 95

RESULT 6
US-09-110-716-31
; Sequence 31, Application US/09110716A
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Glasgow, Benjamin J.
; TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
; FILE REFERENCE: 22000-20596 00
; CURRENT APPLICATION NUMBER: US/09/110,716A
; CURRENT FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Hapophyllin C
;
US-09-110-716-31
```

```

Query Match      87.9%; Score 496; DB 15; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 73
      |||||||
DB 1 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 60

OY 74 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 108
      |||||||
DB 61 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 95

RESULT 7
US-09-412-231-12
; Sequence 12, Application US/09412231
; GENERAL INFORMATION:
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: NOVEL SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: CURA-29 secreted proteins
; CURRENT APPLICATION NUMBER: US/09/412,231
; CURRENT FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: USSN 60/103,195
; PRIOR FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-412-231-12

Query Match      87.9%; Score 496; DB 18; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 73
      |||||||
DB 1 MKLWVLMALALLHCYADSGCKLEDMVEKTI NSDISIPYKELLQEFIDSDAAEAMG 60

OY 74 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 108
      |||||||
DB 61 KFKCFLNOSHRTLNKNGIMHTVYDSIWCNMSN 95

RESULT 8
US-09-471-276-1555
; Sequence 1555, Application US/09471276
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert A.
; TITLE OF INVENTION: EXPRESSED SEQUENCE TAGS AND ENCODED HUMAN PROTEINS.
; FILE REFERENCE: GENSET.025CP1
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent. pm
; SEQ ID NO 1555
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL.
```

LOCATION: -18.-1
US-09-471-276-1555

Query Match 87.9%; Score 496; DB 18; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 108
DB 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 95

RESULT 9
US-09-549-342A-3
Sequence 3, Application US/09549342A
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
APPLICANT: Colpitts, Tracey L.
APPLICANT: Russell, John C.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
FILE REFERENCE: 5972.US.P6
CURRENT APPLICATION NUMBER: US/09/549,342A
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: US 09/467,602
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: US 09/215,818
PRIOR FILING DATE: 1998-12-18
PRIOR APPLICATION NUMBER: US 08/912,276
PRIOR FILING DATE: 1997-08-17
PRIOR APPLICATION NUMBER: US 08/697,105
PRIOR FILING DATE: 1996-08-19
PRIOR APPLICATION NUMBER: US 08/912,149
PRIOR FILING DATE: 1997-08-15
PRIOR APPLICATION NUMBER: US 08/697,106
PRIOR FILING DATE: 1996-08-19
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 95
TYPE: PRT
ORGANISM: Homo sapiens
US-09-549-342A-3

Query Match 87.9%; Score 496; DB 19; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 108
DB 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 95

RESULT 10
US-09-985-911-6
Sequence 6, Application US/09985911
GENERAL INFORMATION:
APPLICANT: NI ET AL.
TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC STEROID-BINDING FACTOR I, II AND III
FILE REFERENCE: PF257D3
CURRENT APPLICATION NUMBER: US/09/985,911
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: 09/583,169

PRIOR FILING DATE: 2000-05-30
PRIOR APPLICATION NUMBER: 09/263,810
PRIOR FILING DATE: 1999-03-08
PRIOR APPLICATION NUMBER: 08/821,451
PRIOR FILING DATE: 1997-03-21
PRIOR APPLICATION NUMBER: 60/014,724
PRIOR FILING DATE: 1996-03-21
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 95
TYPE: PRT
ORGANISM: human
US-09-985-911-6

Query Match 87.9%; Score 496; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 3,4e-49;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 73
DB 1 MKLWVLMALALHCHYADSGCKLLEDVMEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 108
DB 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNKN 95

RESULT 11
US-09-057-719-1075
Sequence 1075, Application US/09057719
GENERAL INFORMATION:
APPLICANT: Edwards, Jean-Baptiste Dumas Milne
APPLICANT: Duclet, Aymeric
TITLE OF INVENTION: EXPRESSED SEQUENCE TAGS FOR SECRETED
NUMBER OF SEQUENCES: 1207
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 West Broadway
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-3505
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,719
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Israelsen, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: GENSET 025A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 1075:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Cancerous prostate
FEATURE:
NAME/KEY: sig-peptide
LOCATION: -18.-1

IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 10.3
OTHER INFORMATION: seq VMLTALLHGYA/DS
US-09-057-719-1075

Query Match 84.2%; Score 475; DB 14; Length 92;
Best Local Similarity 98.9%; Pred. No. 9e-47;
Matches 91; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 14 MKLMTVMTALLHGYADSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMG 73
DB 1 MKLMTVMTALLHGYADSGCKLLEDMVXKTIINSISIPYKELLOEFIDSDAAAEAMG 60

OY 74 KFKOCFLNOSHRTLNKFGLMHMTYDSIMCM 105
DB 61 KFKOCFLNOSHRTLNKFGLMHMTYDSIMCM 92

RESULT 12
US-09-110-716-13
Sequence 13, Application US/09110716A
GENERAL INFORMATION:

APPLICANT: Lehrer, Robert I.
APPLICANT: Zhao, Chengquan
APPLICANT: Glasgow, Benjamin J.
TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
FILE REFERENCE: 22000-20596.00
CURRENT APPLICATION NUMBER: US/09/110.716A
CURRENT FILING DATE: 1998-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 13
LENGTH: 77
TYPE: PRT
ORGANISM: lipophilin C
US-09-110-716-13

Query Match 72.3%; Score 408; DB 15; Length 77;
Best Local Similarity 100.0%; Pred. No. 4.3e-39;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 91
DB 1 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 60

OY 92 LMHMTYDSIMCMKSN 108
DB 61 LMHMTYDSIMCMKSN 77

RESULT 13
US-09-110-716-40
Sequence 40, Application US/09110716A
GENERAL INFORMATION:

APPLICANT: Lehrer, Robert I.
APPLICANT: Zhao, Chengquan
APPLICANT: Glasgow, Benjamin J.
TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
FILE REFERENCE: 22000-20596.00
CURRENT APPLICATION NUMBER: US/09/110.716A
CURRENT FILING DATE: 1998-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 40
LENGTH: 76
TYPE: PRT
ORGANISM: Lpnc
US-09-110-716-40

Query Match 69.8%; Score 393.5; DB 15; Length 76;

Best Local Similarity 98.7%; Pred. No. 2e-37;
Matches 76; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 32 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 91
DB 1 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 60

OY 92 LMHMTYDSIMCMKSN 108
DB 61 LMHMTYDSIMCMKSN 76

RESULT 14
US-09-367-009-3
Sequence 3, Application US/09367009
GENERAL INFORMATION:

APPLICANT: Morris, Carol
APPLICANT: Wilcox, Mark
APPLICANT: Bolis, Shirley
APPLICANT: Walsh, Bradley
APPLICANT: Herbert, Ben
APPLICANT: Molloy, Mark
APPLICANT: Gooley, Andrew Arthur
APPLICANT: Williams, Keith Leslie
TITLE OF INVENTION: Diagnosis of Disease Using Tears
FILE REFERENCE: 47763-5010-US
CURRENT APPLICATION NUMBER: US/09/367.009
CURRENT FILING DATE: 1999-11-08
PRIOR APPLICATION NUMBER: AU P05009
PRIOR FILING DATE: 1997-02-07
PRIOR APPLICATION NUMBER: PCT/AU98/00071
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 74
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Disease marker sequence
US-09-367-009-3

Query Match 69.1%; Score 390; DB 17; Length 74;
Best Local Similarity 98.6%; Pred. No. 5e-37;
Matches 73; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 32 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 91
DB 1 DSGCKLLEDMVEKTIINSISIPYKELLOEFIDSDAAAEAMCKFKOCFLNOSHRTLNKFG 60

OY 92 LMHMTYDSIMCM 105
DB 61 LMHMTYDSIMCM 74

RESULT 15
US-60-234-446-797
Sequence 797, Application US/60234446
GENERAL INFORMATION:

APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
FILE REFERENCE: CLO00832
CURRENT APPLICATION NUMBER: US/60/234.446
CURRENT FILING DATE: 2000-09-18
NUMBER OF SEQ ID NOS: 1797
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 797
LENGTH: 71
TYPE: PRT
ORGANISM: HUMAN

US-60-234-446-797

Query Match 60.88; Score 343; DB 26; Length 71;
 Best Local Similarity 97.14; Pred. No. 1.3e-31;
 Matches 68; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 14 MKLLMVLMLAALLHRCYADSGCKLLEDVVEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 73
 DB 1 MKLLMVLMLAALLHRCYADSGCKLLEDVVEKTIINSDISIPEYKELLOEFIDSDAAAEAMG 60

QY 74 KFKOCFLNOS 83
 DB 61 KFKOCFLNOS 70

Search completed: June 20, 2002, 11:05:26
 Job time: 222 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: June 20, 2002, 11:01:44 ; Search time 12.89 Seconds
(without alignments)
204,652 Million cell updates/sec

Title: US-09-673-395A-238

Perfect score: 564
Sequence: 1 CHARLNTDSSRLAKMLNL.....NFGIMHTYDVISCNKNSN 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	87.9	95	3	US-08-821-451A-6
2	496	87.9	95	4	US-09-263-810-6
3	496	87.9	95	4	US-09-583-169-6
4	281	49.8	93	1	US-08-455-896-2
5	281	49.8	93	2	US-08-933-149-2
6	281	49.8	93	2	US-08-253-2
7	281	49.8	93	3	US-09-082-253-2
8	281	49.8	93	5	PCT-US96-08235-2
9	196.5	34.8	95	3	US-08-821-451A-27
10	196.5	34.8	95	4	US-09-263-810-27
11	196.5	34.8	95	4	US-09-583-169-27
12	195.5	34.7	95	1	US-08-455-896-7
13	195.5	34.7	95	2	US-08-933-149-7
14	195.5	34.7	95	2	US-09-082-253-7
15	195.5	34.7	95	2	US-09-082-253-7
16	195.5	34.7	95	5	PCT-US96-08235-7
17	65	11.5	1255	3	US-08-947-823-3
18	63.5	11.3	582	1	US-08-431-080-16
19	63.5	11.3	582	2	US-08-938-534-16
20	61.5	10.9	1604	2	US-09-004-838-95
21	61	10.8	445	2	US-08-900-148-2
22	61	10.8	446	2	US-08-922-171-3
23	61	10.8	472	2	US-08-922-171-2
24	60.5	10.7	1257	3	US-08-947-823-5
25	60.5	10.7	971	3	US-09-112-450-2
26	60.5	10.7	2471	3	US-09-112-450-4
27	59.5	10.5	114	1	US-08-031-399-3

28	59.5	10.5	114	1	US-08-031-399-6	Sequence 6, Appl1
29	59.5	10.5	114	1	US-08-031-399-12	Sequence 12, Appl1
30	59.5	10.5	114	1	US-08-393-305-3	Sequence 3, Appl1
31	59.5	10.5	114	1	US-08-393-305-6	Sequence 6, Appl1
32	59.5	10.5	114	1	US-08-726-817-3	Sequence 3, Appl1
33	59.5	10.5	114	1	US-08-726-817-6	Sequence 6, Appl1
34	59.5	10.5	114	1	US-08-504-042-3	Sequence 3, Appl1
35	59.5	10.5	114	1	US-08-504-042-6	Sequence 6, Appl1
36	59.5	10.5	114	1	US-08-504-042-12	Sequence 12, Appl1
37	59.5	10.5	114	2	US-08-725-969-3	Sequence 3, Appl1
38	59.5	10.5	114	2	US-08-725-969-6	Sequence 6, Appl1
39	59.5	10.5	114	2	US-08-794-524-3	Sequence 3, Appl1
40	59.5	10.5	114	2	US-08-794-524-6	Sequence 6, Appl1
41	59.5	10.5	114	4	US-09-189-193-3	Sequence 3, Appl1
42	59.5	10.5	114	4	US-09-189-193-6	Sequence 6, Appl1
43	59.5	10.5	114	5	PCT-US94-03793-3	Sequence 3, Appl1
44	59.5	10.5	114	5	PCT-US94-03793-6	Sequence 6, Appl1
45	59.5	10.5	114	5	PCT-US94-03793-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-08-821-451A-6
; Sequence 6, Application US/08821451A
; Patent No. 6066724
; GENERAL INFORMATION:
; APPLICANT: Jian NL, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CECCHI, STEFANO, BAIN, GILFILLAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,451A
; FILING DATE: March 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,724
; FILING DATE: March 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (P257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SPO ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-821-451A-6

Query Match 87.9%; Score 496; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 8,8e-54;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	14	MLKLMTMTALAILLHCTYAASGCKLLEDVETINSIDSIPEYKELQEFIDSDAAAEAAAG	73
Db	1	MLKLMTMTALAILLHCTYAASGCKLLEDVETINSIDSIPEYKELQEFIDSDAAAEAAAG	60
Oy	74	KFKOCFLNOSHRTLRKEGLMTHTVYDSITMCNKNKSN	108
Db	61	KFKOCFLNOSHRTLRKNEGLMTHTVYDSITMCNKNKSN	95

```

1      RESULT      2
2      US-09-263-810-6
3      : Sequence 6, Application US/09263810
4      : Patent No. 6174992
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
9      :
10     : TITLE OF INVENTION: Human Endometrial Specific Steroid-
11     :
12     : TITLE OF INVENTION: Binding Factor I, II and III
13     :
14     : NUMBER OF SEQUENCES: 27
15     :
16     : CORRESPONDENCE ADDRESS:
17     :
18     : ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
19     :
20     : ADDRESSEE: CECCHI, STEWART & OLSTEIN
21     :
22     : STREET: 6 BECKER FARM ROAD
23     :
24     : CITY: ROSELAND
25     :
26     : STATE: NEW JERSEY
27     :
28     : COUNTRY: USA
29     :
30     : ZIP: 07068
31     :
32     : COMPUTER READABLE FORM:
33     :
34     : MEDIUM TYPE: 3.5 INCH DISKETTE
35     :
36     : COMPUTER: IBM PS/2
37     :
38     : OPERATING SYSTEM: MS-DOS
39     :
40     : SOFTWARE: WORD PERFECT 5.1
41     :
42     : CURRENT APPLICATION DATA:
43     :
44     : APPLICATION NUMBER: US/09/263,810
45     :
46     : FILING DATE:
47     :
48     : CLASSIFICATION:
49     :
50     : PRIOR APPLICATION DATA:
51     :
52     : APPLICATION NUMBER: 08/821,451
53     :
54     : FILING DATE:
55     :
56     : ATTORNEY/AGENT INFORMATION:
57     :
58     : NAME: MULILINS, J. G.
59     :
60     : REGISTRATION NUMBER: 33,073
61     :
62     : REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
63     :
64     : TELECOMMUNICATION INFORMATION:
65     :
66     : TELEPHONE: 201-994-1700
67     :
68     : TELEFAX: 201-994-1744
69     :
70     : INFORMATION FOR SEQ ID NO: 6:
71     :
72     : SEQUENCE CHARACTERISTICS:
73     :
74     : LENGTH: 95 AMINO ACIDS
75     :
76     : TYPE: AMINO ACID
77     :
78     : STRANDEDNESS:
79     :
80     : TOPOLOGY: LINEAR
81     :
82     : MOLECULE TYPE: PROTEIN
83     :
84     :
85     : US-09-263-810-6

```

```

Query Match Similarity      87.9% ; Score 496 ; DB 4 ; Length 95 ;
Best Local Similarity      100.0% ; Pred. No. 8,8e-54 ;
Matches      95 ; Conservative      0 ; Mismatches      0 ; Indels      0 ; Gaps      0 ;

OY      14 MKLLVYMLAAILLHCYADSGCKLEDMYEKTIINSISIPERKELLOEITDSDAAAEANG      73
          |||
Db       1 MKLLVYMLAAILLHCYADSGCKLEDMYEKTIINSISIPERKELLOEITDSDAAAEANG      60

OY      74 KFKQCFLNQSHPTLKNFGMLMHTVYDSIWCNKNKS      108
          |||
Db       61 KFKQCFLNQSHPTLKNFGMLMHTVYDSIWCNKNKS      95

RESULT      3
US-09-583-169-6
; Sequence 6, Application US/09583169
; Patent No. 6338948
; GENERAL INFORMATION:

```

```

1  APPLICANT : Jian NI, Guo-Liang Yu and Reiter Genltz
2  TITLE OF INVENTION : Human Endometrial Specific Steroid-
3  TITLE OF INVENTION : Binding Factor I, II and III
4  NUMBER OF SEQUENCES: 27
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE: CARELLA, BYRNE, BAIN, GIFFILLAN,
7  ADDRESSEE: CECCHI, STEWART & OLSTEIN
8  STREET: 6 BECKER FARM ROAD
9  CITY: ROSELAND
10 STATE: NEW JERSEY
11 COUNTRY: USA
12 ZIP: 07068
13
14 COMPUTER READABLE FORM:
15 MEDIUM TYPE: 3.5 INCH DISKETTE
16
17 COMPUTER: IBM PS/2
18 OPERATING SYSTEM: MS-DOS
19 SOFTWARE: WORD PERFECT 5.1
20
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/09/583,169
23
24 FILING DATE:
25 CLASSIFICATION:
26
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: 08/821,451
29
30 FILING DATE:
31 ATTORNEY/AGENT INFORMATION:
32 NAME: MULLINS, J.G.
33 REGISTRATION NUMBER: 33,073
34 REFERENCE/DOCKET NUMBER: 325800-521 (P257)
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: 201-994-1700
37 TELEFAX: 201-994-1744
38
39 INFORMATION FOR SEQ. ID NO.: 6:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 95 AMINO ACIDS
42 TYPE: AMINO ACID
43 STRANDEDNESS:
44 TOPOLOGY: LINEAR
45
46 MOLECULE TYPE: PROTEIN
47
48 US-09-583-169-6

```

```

0 Query Match Similarity      87.9%: Score 496; DB 4; Length 95;
1 Best Local Similarity      100.0%; Pred. No. 8.8e-54;
2 Matches    95; Conservative    0; Mismatches    0; Indels    0; Gaps    0
3
4
5 RESULT          4
6 US-08-455-896-2
7 : Sequence 2, Application US/08455896
8 : Patent No. 5668267
9 : GENERAL INFORMATION:
10 : APPLICANT: WATSON, MARK A.
11 : TITLE OF INVENTION: FLEMING, TIMOTHY P.
12 : TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
13 : NUMBER OF SEQUENCES: 13
14 : CORRESPONDENCE ADDRESS:
15 : ADDRESSEE: ROGERS, HOWELL & HAEFERKAMP
16 : STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
17 : CITY: ST. LOUIS
18 : STATE: MISSOURI
19 : COUNTRY: USA
20 : ZIP: 63105-1817
21 : COMPUTER READABLE FORM:
22 : MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,896
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 952726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-455-896-2

Query Match 49.8%; Score 281; DB 1; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,4e-27;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Caps 0;

QY 14 MKLWVLMALALLHCYAGSCCKLEDEWEKTIINSDISIPYKELLOFIDSDAAAEAMG 73
DB 1 MKLWVLMALASQHCYAGSCCKPLENVIKTIINQVSKTEYKELLOFIDDMATTNAID 60
QY 74 KFKOCFLNOSHRTLNKFGIMHMYVDSIWCNM 105
DB 61 ELKECFLNQTDFTLSNVEVFMOIYDSSLCDL 92

RESULT 5
US-08-933-149-2
Sequence 2, Application US/08933149
Patent No. 5922836
GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: MAMMAGLOBIN, A SECRETED
TITLE OF INVENTION: MAMMARY SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,149
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: HENDERSON, MELODIE W.
REGISTRATION NUMBER: 37,848
REFERENCE/DOCKET NUMBER: 6029-6040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-933-149-2

Query Match 49.8%; Score 281; DB 2; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,4e-27;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Caps 0;

QY 14 MKLWVLMALALLHCYAGSCCKLEDEWEKTIINSDISIPYKELLOFIDSDAAAEAMG 73
DB 1 MKLWVLMALASQHCYAGSCCKPLENVIKTIINQVSKTEYKELLOFIDDMATTNAID 60
QY 74 KFKOCFLNOSHRTLNKFGIMHMYVDSIWCNM 105
DB 61 ELKECFLNQTDFTLSNVEVFMOIYDSSLCDL 92

RESULT 6
US-09-082-343-2
Sequence 2, Application US/09082343
Patent No. 5968754
GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,343
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/455,896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 952726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-082-343-2

Query Match 49.8%; Score 281; DB 2; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,4e-27;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Caps 0;

```

OY      14  MKLIMVMTLALLLHCYAAAGCKLLLEDVAKTINSDSIPEYKELLOEFPIDSAALAEAG  73
Db      1  MKLIMVMTLALLLHCYAAAGCGPLELANTISTINQVSKTEYKELLOEFPIDNATTAID  60
OY      74  KFKOCFLNOSHRLTKNFGMLMTIYVDSIMCN  105
Db      61  ELKECFPLNODTEFLSNVEFMOLITDSSLCOL  92

```

```

1      RESULT      7
2      US-09-082-253-2
3      : Sequence 2, Application US/09082253
4      : Patent No. 6004756
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: WATSON, MARK A.
9      : APPLICANT: FLEMING, TIMOTHY P.
10     : TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
11     : TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
12     : NUMBER OF SEQUENCES: 13
13     : CORRESPONDENCE ADDRESS:
14     : ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
15     : STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
16     : CITY: ST. LOUIS
17     : STATE: MISSOURI
18     : COUNTRY: USA
19     : ZIP: 63105-1817
20     :
21     : COMPUTER READABLE FORM:
22     : MEDIUM TYPE: Floppy disk
23     : COMPUTER: IBM PC compatible
24     : OPERATING SYSTEM: PC-DOS/MS-DOS
25     : SOFTWARE: PatentIn Release #1.0, Version #1.25
26     :
27     : CURRENT APPLICATION DATA:
28     : APPLICATION NUMBER: US/09/082,253
29     : FILING DATE:
30     : CLASSIFICATION:
31     :
32     : PRIOR APPLICATION DATA:
33     : APPLICATION NUMBER: 08/455,896
34     : FILING DATE: 05/31/1995
35     : ATTORNEY/AGENT INFORMATION:
36     : NAME: HOLLAND, DONALD R.
37     : REGISTRATION NUMBER: 35,197
38     : REFERENCE/DOCKET NUMBER: 953726
39     : TELECOMMUNICATION INFORMATION:
40     : TELEPHONE: (314) 727-5188
41     : TELEFAX: (314) 727-6092
42     :
43     : INFORMATION FOR SEQ ID NO: 2:
44     : SEQUENCE CHARACTERISTICS:
45     : LENGTH: 93 amino acids
46     : TYPE: amino acid
47     : STRANDEDNESS: single
48     : TOPOLOGY: linear
49     : MOLECULE TYPE: protein
50     :
51     : HYPOTHETICAL: NO
52     :
53     : US-09-082-253-2

```

```

Query Match Similarity      49.8%; Score 281; DB 3; Length 93;
Best Local Similarity      58.7%; Pred. No. 2,4e+27;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY      14 MKLIMVLMIALLLHCYADSGCKLEDVMEKTI NSDI SIPEYKELIQEITDSDAAEAMC 73
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MKLIMVLMIALLSQHCYAGSGCPLEENVIKTI NSPVSSTKEVKELOEITDNNATNAID 60
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY      74 KKKQCFILNOSHRTLKNFGMLMHTVYDSINCM 105
          : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61 ELKECFILNQTDETLINVEYFMQLITDSSICDL 92
          : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT      8
PCT-US96--08235-2
; Sequence 2, Application PC/TUS9608235
; GENERAL INFORMATION:

```

APPLICANT: MATSON, MARK A.
 APPLICANT: FLEMING, TIMOTHY P.
 TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
 TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ROGERS, HOWELL, & HAFERKAMP
 STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
 CITY: ST. LOUIS
 STATE: MISSOURI
 COUNTRY: USA
 ZIP: 63105-1817
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/08235
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: HOLLAND, DONALD R.
 REGISTRATION NUMBER: 35,197
 REFERENCE/DOCKET NUMBER: 964796
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (314) 727-5188
 TELEFAX: (314) 727-6092
 INFORMATION FOR SEQ. ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 93 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEetical: NO

Query Match	49.8%;	Score 281;	DB 5;	Length 93;
Best Local Similarity	58.7%;	Pred. No. 2,48-27;		
Matches	54;	Conservative 13;	Mismatches 25;	Indels 0;
			Gaps	0;
QY	14	MLKLMLMLALALHCAVDSGCKLLIDYENKRTINSISIPKRYKELLOFIDSDPAARAC	73	
DB	1	MLKLMLMLALALSHCAAGSCPELENYSKRTINPQVSKTEKKELOFEDIDNATTAID	60	
QY	74	KFKQCFILNQSHRTILKNFGLMMHTVYSINCMN	105	
DB	1	ELKECFILNQSHRTILSNVFFQGLITDSICDIL	92	

RESULT 9
 US-08-821-451A-27
 : Sequence 27, Application US/08821451A
 : Patent No. 6066724
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
 :
 : TITLE OF INVENTION: Human Endometrial Specific Steroid
 :
 : TITLE OF INVENTION: Binding Factor 1, II and III
 :
 : NUMBER OF SEQUENCES: 27
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: CARBELL, BYRNE, BAIN, GILFILLAN,
 :
 : ADDRESSEE: CECCHI, STEWART & OLSTEIN
 :
 : STREET: 6 BECKER FARM ROAD
 :
 : CITY: ROSELAND
 :
 : STATE: NEW JERSEY
 :
 : COUNTRY: USA
 :
 : ZIP: 07068
 :
 : COMPUTER READABLE FORM:
 :
 : MEDIUM TYPE: 3.5 INCH DISKETTE
 :
 : COMPUTER: IBM PS/2
 :
 : OPERATING SYSTEM: MS-DOS


```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,343
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/455,896
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188
; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-09-082-343-7

```

```

Query Match          34.7%; Score 195.5; DB 2; Length 95;
Best Local Similarity 35.8%; Pred. No. 8e-17;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1;

```

```

QY 14 MKLLVLMALAILHGYAD-SGCKLLEDVKEKTSIPYKELQEFIDSDAAAEAM 72
   |||: : : : : ||| ||| : : : : : ||| : : : : : : : : : : : :
DB 1 MKLVFLFLVLTIPICCYASGSGSILDEVIRGTINSTVTLHDYMKLVKPYVDHFTKAV 60
   : |||||: : : : : ||| : : : : : : : : : : : :
QY 73 GKFKOCFLNOSHRTLKNFGIMHTVYDSIWCNKS 107
   : |||||: : : : : ||| : : : : : : : : : : : :
DB 61 KQFKOCFLDQTKLTLENVGVMEALFNSESCQOPS 95

```

```

RESULT 15
US-09-082-253-7
; Sequence 7, Application US/09082253
; Patent No. 6004756
; GENERAL INFORMATION:
; APPLICANT: WATSON, MARK A.
; APPLICANT: FLEMING, TIMOTHY P.
; TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
; TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ROGERS, HOWELL & HAFERRAMP
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,253
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/455,896
; FILING DATE: 05/31/1995
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 952726
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314) 727-5188

```

```

; TELEFAX: (314) 727-6092
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; US-09-082-253-7

```

```

Query Match          34.7%; Score 195.5; DB 3; Length 95;
Best Local Similarity 35.8%; Pred. No. 8e-17;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1;

```

```

QY 14 MKLLVLMALAILHGYAD-SGCKLLEDVKEKTSIPYKELQEFIDSDAAAEAM 72
   |||: : : : : ||| ||| : : : : : ||| : : : : : : : : : : : :
DB 1 MKLVFLFLVLTIPICCYASGSGSILDEVIRGTINSTVTLHDYMKLVKPYVDHFTKAV 60
   : |||||: : : : : ||| : : : : : : : : : : : :
QY 73 GKFKOCFLNOSHRTLKNFGIMHTVYDSIWCNKS 107
   : |||||: : : : : ||| : : : : : : : : : : : :
DB 61 KQFKOCFLDQTKLTLENVGVMEALFNSESCQOPS 95

```

```

Search completed: June 20, 2002, 11:02:06
Job time: 22 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2002, 11:01:44 ; Search time 30.14 Seconds
(without alignments)
398.008 Million cell updates/sec

Title: US-09-673-395A-238
Perfect score: 564
Sequence: 1 CHARLNTDSSRLANKLMLV.....NFGIMHTVYDSTICWKNKS 108

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_032802:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	564	100.0	108	20	AA60038 Human endometrium
2	496	87.9	95	18	AA60038 Human endometrium
3	496	87.9	95	20	AA60038 Human endometrium
4	496	87.9	95	21	AA60038 Human endometrium
5	496	87.9	95	21	AA60038 Human endometrium
6	496	87.9	95	21	AA60038 Human endometrium
7	496	87.9	95	21	AA60038 Human endometrium
8	496	87.9	95	22	AA60038 Human endometrium
9	390	69.1	74	19	AA60038 Human endometrium
10	281	49.8	93	18	AA60038 Human endometrium
11	281	49.8	93	19	AA60038 Human endometrium

12	281	49.8	93	19	AA60038 Human endometrium
13	281	49.8	93	20	AA60038 Human endometrium
14	281	49.8	93	21	AA60038 Human endometrium
15	281	49.8	93	21	AA60038 Human endometrium
16	281	49.8	93	22	AA60038 Human endometrium
17	281	49.8	93	22	AA60038 Human endometrium
18	281	49.8	93	22	AA60038 Human endometrium
19	281	49.8	93	22	AA60038 Human endometrium
20	281	49.8	93	22	AA60038 Human endometrium
21	281	49.8	93	22	AA60038 Human endometrium
22	281	49.8	93	22	AA60038 Human endometrium
23	277	49.1	93	22	AA60038 Human endometrium
24	277	49.1	93	22	AA60038 Human endometrium
25	277	49.1	93	22	AA60038 Human endometrium
26	273	48.4	93	22	AA60038 Human endometrium
27	272	48.2	93	22	AA60038 Human endometrium
28	269	47.7	93	22	AA60038 Human endometrium
29	265.5	47.1	90	22	AA60038 Human endometrium
30	265.5	47.1	90	22	AA60038 Human endometrium
31	206	36.5	74	21	AA60038 Human endometrium
32	95	16.8	33	19	AA60038 Human endometrium
33	77	13.7	21	22	AA60038 Human endometrium
34	77	13.7	21	22	AA60038 Human endometrium
35	77	13.7	21	22	AA60038 Human endometrium
36	77	13.7	21	22	AA60038 Human endometrium
37	73	12.9	702	22	AA60038 Human endometrium
38	72	12.6	20	22	AA60038 Human endometrium
39	71	12.6	828	22	AA60038 Human endometrium
40	71	12.6	1187	22	AA60038 Human endometrium
41	70.5	12.5	343	22	AA60038 Human endometrium
42	70	12.4	457	22	AA60038 Human endometrium
43	68.5	12.1	384	22	AA60038 Human endometrium
44	68.5	12.1	619	22	AA60038 Human endometrium
45	68	12.1	55	22	AA60038 Human endometrium

ALIGNMENTS

RESULT 1	AA60038	standard: protein; 108 AA.
ID	AA60038	standard: protein; 108 AA.
AC	AA60038	
XX	31-JAN-2000	(first entry)
DT	Human endometrium	tumour EST encoded protein 98.
DE	Human endometrium	tumour; cancer; anticancer; cytostatic; EST:
KW	treatment; uterine; gene therapy; expressed sequence tag.	
KW	Homo sapiens.	
OS	DE19817948-A1.	
PN	21-OCT-1999.	
PD	17-APR-1998;	98DE-1017948.
PF	17-APR-1998;	98DE-1017948.
PR	17-APR-1998;	98DE-1017948.
XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.	
PA	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;	
PI	WPI: 1999-591957/51.	
XX	N-PSDB; AA42013.	
DR	New nucleic acid sequences expressed in uterine cancer tissues, and	
XX	derived polypeptides, for treatment of uterine and endometrial cancer	
PT	and identification of therapeutic agents	
XX		

PS Claim 23: Page 314: 444pp; German.

XX This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer; (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for
 CC generation of specific antibodies. (A) are identified by assembling ESTs
 CC (expressed sequence tags) from a particular tissue type before comparison
 CC of expression patterns. This allows a significantly longer fragment of
 CC the gene to be revealed, so should reduce the number of failures may represent
 CC associated with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. AAY59941-Y60328 represent
 CC protein fragments encoded by the human endometrium tumour cDNA library
 CC derived EST fragments represented in AA41981-442121.

SO Sequence 108 AA:

Query Match 100.0%; Score 564; DB 20; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.5e-53; Indels 0; Gaps 0;
 Matches 108; Conservative 0; Mismatches 0;

OY 1 CHARLNTDSRLAKMLVLMALALHCVADSGCKLEDMVEKINSISIPYKELIQ 60

Db 1 charlntdsrlakmlvmlaalhhcyadsgcklleedmvktinsdisipykellq 60

OY 61 EFIDSDAAAEAMGKFKOCFLNOSHTLKNFGIMHTYDSTICMKN 108

Db 61 efidsdaaeamgkfkqclnshltknfgimhtydsicmkn 108

RESULT 2

AAW35804 standard; Protein: 95 AA.

AAW35804:

27-MAR-1998 (first entry)

Human endometrial specific steroid-binding factor III.

Endometrial specific steroid-binding factor III; ESR III; human;

Clara cell secretory protein; endometrium;

phospholipase A2 inhibitor; polychlorinated biphenyl; antiagregant;

Inflammation; asthma; rhinitis; cystic fibrosis; airway disease;

neoplasia; atopy; therapy; diagnosis.

Homo sapiens.

Key Location/Qualifiers

Peptide 1..21

Protein /label= Sig_peptide

25-SEP-1997.

21-MAR-1996; 96WO-US03857.

21-MAR-1996; 96WO-US03857.

(HUMA-) HUMAN GENOME SCI INC.

Gentz RL, Ni J, Yu G;

WPI: 1997-480206/44.

DR N-PSDB: AAT94832.

XX Human endometrial specific steroid-binding factor I, II and III -
 PT used to treat inflammation, asthma, rhinitis, cystic fibrosis,
 PT airway disease, neoplasia, atopy etc.

PS Claim 19: Page 65; 92pp; English.

XX This sequence comprises human endometrial specific steroid binding
 CC factor III (ESF III), a protein that inhibits phospholipase A2
 CC activity, binds to polychlorinated biphenyl compounds, reduces
 CC foreign protein antigenicity, inhibits monocyte and neutrophil
 CC chemotaxis and phagocytosis, inhibits platelet aggregation, the
 CC regulates eicosanoid levels in the human uterus and controls the
 CC growth of endometrial cells. The amino acid sequence was deduced
 CC from a cDNA clone (see AAT94832) derived from a human endometrial
 CC tumour. ESR I (see AAW35802) and ESR II (see AAW35803) are also
 CC claimed. Human ESR III has about 36% identity with rat prostatic
 CC steroid-binding protein. Recombinant ESR I, II and III can be
 CC expressed in host cells for use in claimed methods (a) for treating
 CC a patient in need of ESR I, II or III (including compounds which bind
 CC to and inhibit activation of the ESR polypeptide, hESF I, II and
 CC III may be used to treat inflammation, asthma, rhinitis, cystic
 CC fibrosis, airway disease, neoplasia and atopy.

SO Sequence 95 AA:

Query Match 87.9%; Score 496; DB 18; Length 95;
 Best Local Similarity 100.0%; Pred. No. 2.9e-46; Indels 0; Gaps 0;
 Matches 95; Conservative 0; Mismatches 0;

OY 14 MKLIMVLMALALHCVADSGCKLEDMVEKINSISIPYKELIQEFIDSDAAEAMG 73

Db 1 mklmvlmalalhhcyadsgcklleedmvktinsdisipykellqefidsdaaeamg 73

OY 74 KFKOCFLNOSHTLKNFGIMHTYDSTICMKN 108

Db 61 kfkqclnshltknfgimhtydsicmkn 95

RESULT 3

AAW02590 standard; Protein: 95 AA.

AAW02590:

26-JUL-1999 (first entry)

A human mamnoglobulin homologue (HMH).

Human mamnoglobulin homologue; HMH; antagonist; neoplastic disorder;

adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;

teratocarcinoma; endometrios.

Homo sapiens.

WO9919487-A1.

22-APR-1999.

14-OCT-1998; 98WO-US21729.

16-OCT-1997; 97US-0951750.

(INCY-) INCYTE PHARM INC.

Hillman JL, Murry LE, Shah P;

WPI: 1999-302531/25.

DR N-PSDB: AAX36138.

PT New human mamnoglobulin homologue (HMH), useful for diagnosing, treating

PT Diagnosing, staging and monitoring gynecological cancer comprising
 PT using an elevated level of ESBP111 in a patient as an indicator of
 PT cancer

PS Claim 6; Page 28-29; 32pp; English.

CC The levels of human endometrial specific steroid binding factor
 CC (ESBP111) can be measured and compared to control levels and used to
 CC diagnose the presence of a gynaecological (uterine, breast, endometrial,
 CC or ovarian) cancer in a patient. ESBP111 levels can also be used to
 CC diagnose metastasis, to stage or monitor gynaecological cancer.
 CC Antibodies specific for ESBP111 can be used to treat gynecological
 CC cancers.

XX Sequence 95 AA:

Query Match 87.9%; Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 2.9e-46;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWMLAALLHCYADSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAEAMG 73
 Db 1 mklmwmlaalllhcysdsgcklledmvektinsdisipeykellqefidsdaaamg 60

OY 74 KFKQCFINQSHRTLNFGIMHFTVYDSIMCNMKN 108
 Db 61 kfkqcfinsgshrtlnfgimhftvdsiwcnmkn 95

RESULT 6

AAV92237
 ID AAV92237 standard; Protein; 95 AA.

AC AAV92237;

DT 10-AUG-2000 (first entry)

DE Mammoglobin homologue from clone Mamm-X.

XX Clone Mamm-X; mammoglobin; breast cancer; cytostatic; anti-HIV;
 KW immunosuppressive; antiarteriosclerotic; antiinflammatory;
 KW antiarteritic; antiarteriosclerotic; vasotrophic; neuroprotective;
 KW nootropic; dermatological; tranquilizer; vulnery.

OS Homo sapiens.

PN WO200020447-A2.

PD 13-APR-2000.

PF 06-OCT-1999; 99WO-US22294.

PR 06-OCT-1998; 98US-0103195.

PR 05-OCT-1999; 99US-0103195.

PA (CURA-) CURAGEN CORP.

PI Shinkets RA;

DR WPI; 2000-303741/26.

DR N-PSDB; AAA09118.

PT Nucleic acids encoding polypeptides with syncollin-like, claudin-like or
 PT cytochrome-like activity, useful for treating diseases including cancer,
 PT Alzheimer's and atherosclerosis

PS Claim 23; Fig 6; 118pp; English.

CC Clone Mamm-X encodes a polypeptide that is 100 percent identical to human
 CC Mammoglobin B precursor, a potential marker of breast cancer nodal
 CC metastasis. The sequences are useful for treatment of diseases such as
 CC cancer, immune disorders, autoimmune disease, acquired immune deficiency

CC syndrome (AIDS), transplant rejection, allergy, infection by a
 CC pathological agent or organism, inflammatory disorders, arthritis, a
 CC haematopoietic disorder, a skin disorder, atherosclerosis, restenosis, a
 CC neurological disease, Alzheimer's disease, trauma, spinal cord injury
 CC and skeletal disorders.

XX Sequence 95 AA:

Query Match 87.9%; Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 2.9e-46;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 MKLWMLAALLHCYADSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAEAMG 73
 Db 1 mklmwmlaalllhcysdsgcklledmvektinsdisipeykellqefidsdaaamg 60

OY 74 KFKQCFINQSHRTLNFGIMHFTVYDSIMCNMKN 108
 Db 61 kfkqcfinsgshrtlnfgimhftvdsiwcnmkn 95

RESULT 7

AAV65394
 ID AAV65394 standard; Protein; 95 AA.

AC AAV65394;

DT 01-FEB-2000 (first entry)

DE Human 5' EST related polypeptide SEQ ID NO:1555.

XX Human; 5' EST; expressed sequence tag; secreted protein; diagnosis;
 KW gene therapy; chromosome mapping; upstream regulatory sequence;
 KW forensic; location; development; protein synthesis; stability;
 KW regulation; identification.

OS Homo sapiens.

PN WO9953051-A2.

PD 21-OCT-1999.

PF 09-APR-1999; 99WO-1B00712.

PR 09-APR-1998; 98US-0057719.

PR 28-APR-1998; 98US-0069047.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-038446/03.

DR N-PSDB; AA243008.

PT Novel secreted protein 5' expressed sequence tag sequences used in
 PT diagnostic, forensic, gene therapy, and chromosome mapping procedures

PS Claim 3; Page 818; 837pp; English.

CC AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
 CC sequences, corresponding to human secreted proteins. AA164651 to
 CC AA165438 represent the EST-related proteins corresponding to AA242265 to
 CC AA243052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be

Matches 73: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 32 DSGCKLEDMVEKTSIDSIPEYKELQEFIDSDAAAPAKRFKOCFLNOSHRLTKNG 91
 1 dsqcklledmvektlndisidipeykellqefidsdaaamgkfkycflnshrllknlg 60

OY 92 LMMHTVYDSICNM 105
 61 lmmhtvydsicnm 74

RESULT 10
 AAW10179
 ID AAW10179 standard; Protein; 93 AA.

XX AAW10179;
 AC
 XX 12-AUG-1997 (first entry)
 DT
 XX
 DE Mammary-specific secretory protein, mammagloblin.

XX
 KW mammagloblin; mammary-secretory protein; breast cancer; detection;
 KW neoplastic disease; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..19 /label= signal_peptide

FT Protein 20..93 /label= mature_protein

XX WO9638463-A1.

XX 05-DEC-1996.

XX 31-MAY-1996; 96WO-US08235.

XX 31-MAY-1995; 95US-0455896.

XX (UNIV) UNIV WASHINGTON.

XX Fleming TP, Watson MA;

XX WPI; 1997-034299/03.

XX N-PSDB; AAT50925.

XX Nucleic acid encoding mammary-specific secretory protein,
 PT mammagloblin - used to develop prods. for the early diagnosis and
 PT treatment of breast cancer neoplastic disease

XX Claim 3; Fig 2; 54pp; English.

XX The present sequence is that of a mammary-specific secretory protein
 CC designated mammagloblin, which is overexpressed in 27% of stage I primary
 CC breast cancer tumours. The anonymous sequence tag previously designated
 CC DSGR002 was used to demonstrate that mammagloblin is abundant in the
 CC breast cancer tumour cell line MDA-MB-415. To isolate the full-length
 CC mammagloblin cDNA (AAT50925), the mRNA was reverse transcribed from
 CC this cell line and cloned using the RACE PCR technique. The nucleic acid
 CC and protein can be used to develop prods. e.g. antibodies or probes, for
 CC the detection and treatment of breast neoplastic disease.

XX Sequence 93 AA;

Query Match 49.8%; Score 281; DB 18; Length 93;
 Best Local Similarity 58.7%; Pred. No. 5.7e-23;
 Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

14 MKLIMVIMLALLLHCTADSGCKLEDMVEKTSIDSIPEYKELQEFIDSDAAAPAK 73
 14 :|||||: ||| ||| |||: ||| :| ||||| ||||| :| :|
 mklmvlmalaalshqcyagsgcpllenvisktlmpyvkelykellqefidnatnaid 60

OY 74 KFKOCFLNOSHRLTKNGIAMHTVYDSICNM 105
 61 elkecfllngtdetlsnvevfmglydsslcl 92

RESULT 11
 AAW59777
 ID AAW59777 standard; Protein; 93 AA.

XX AAW59777;
 AC
 XX 12-OCT-1998 (first entry)
 DT
 XX
 DE Amino acid sequence of the human steroid binding protein C2.

XX
 KW Human steroid-binding protein C2; hsbp2; hsbp1; breast cancer; probe;
 KW gene therapy vector; ribozyme; probe; hybridisation; amplification;
 KW antibody; immunoassay.

XX Homo sapiens.
 OS
 XX WO9821331-A1.
 PN
 XX 22-MAY-1998.
 PD
 XX 07-NOV-1997; 97WO-US20674.
 PF
 XX 12-NOV-1996; 96US-0747547.
 PR
 XX (INCY-) INCYTE PHARM INC.

XX Akerblom IE, Goli SK, Hawkins PR, Hillman JL, Murry LE;
 PI
 XX WPI; 1998-297935/26.
 DR
 XX N-PSDB; AAV41580.

XX New human steroid binding proteins C1 and C2 - useful for, e.g.
 PT diagnosis, monitoring and treating breast cancer, and for drug
 PT screening

XX Claim 12; Fig 2; 70pp; English.

XX This is the amino acid sequence of the human steroid-binding protein
 CC C2 (hsbp2) used in the method of the invention for the diagnosis,
 CC monitoring and treatment of breast cancer. Hsbp1 and hsbp2 are useful
 CC as markers for breast cancer, i.e. measuring levels of hsbp1 and hsbp2
 CC used for diagnosis or monitoring the disease, to identify subjects
 CC at risk and to discriminate between different forms of cancer for
 CC selection of appropriate therapies. They may also be used for drug
 CC screening. Nucleic acids encoding hsbp1 and hsbp2 can be used in gene
 CC therapy vectors to over express the steroid-binding proteins, preventing
 CC binding of steroids, or antisense sequences, ribozymes. Their nucleic
 CC acids can also be used for the diagnosis and monitoring (by quantifying
 CC expression of hsbp), as source of probes for hybridisation and
 CC amplification of genomic or related sequences for studying regulation of
 CC gene function and for mapping the genomic sequence. Antibodies are used
 CC as diagnostic reagents in standard immunoassays for hsbp.

XX Sequence 93 AA;

Query Match 49.8%; Score 281; DB 19; Length 93;
 Best Local Similarity 58.7%; Pred. No. 5.7e-23;
 Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 14 MKLIMVIMLALLLHCTADSGCKLEDMVEKTSIDSIPEYKELQEFIDSDAAAPAK 73
 14 :|||||: ||| ||| |||: ||| :| ||||| ||||| :| :|
 Db 1 mklmvlmalaalshqcyagsgcpllenvisktlmpyvkelykellqefidnatnaid 60

OY 74 KFKOCFLNOSHRLTKNGCLMMHTVYDSICNM 105
 74 :|||||: ||| ||| |||: ||| :| ||||| ||||| :| :|
 Db 61 elkecfllngtdetlsnvevfmglydsslcl 92

RESULT 12

AAW48432
ID AAW48432 standard; Protein; 93 AA.

AC AAW48432;

DT 13-JUL-1998 (first entry)

XX Mammaglobin protein.

DE Mammaglobin; detection; diagnosis; breast cancer; tumour; antibody;

XX gene therapy; human.

KW Homo sapiens.

OS WO9807753-A1.

PN 26-FEB-1998.

PD 19-AUG-1997; 97WO-US14666.

PF 15-AUG-1997; 97US-0697106.

PR 19-AUG-1996; 96US-0697106.

XX (ABBO) ABBOTT LAB.

PA Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;

PI Roberts-Rapp L, Russell JC, Stroupe SD;

DR WPI; 1998-169096/15.

DR N-PSDB; AAV17905, AAV17906.

XX Antibodies to mammaglobin polypeptide(s) - used for detecting,

PT diagnosing, preventing or treating diseases or conditions of breast

PT such as breast cancer

XX Claim 8; Page 92; 105pp; English.

The present sequence represents mammaglobin which is used in an example of the present invention. The present invention describes an antibody (A) which specifically binds to at least 1 mammaglobin epitope (ME) which is derived from an amino acid sequence having at least 50% identity to an amino acid sequence (see AAW48432) and fragments. Also described are: (1) an assay kit for determining the presence of mammaglobin antigen (MA) in a test sample, comprising a container containing an antibody as in (A); (2) a method for producing antibodies which specifically bind to a MA, comprising administering an isolated immunogenic polypeptide or fragment to elicit an immune response, where the immunogenic polypeptide comprises at least 1 ME and has at least 50% identity to a sequence (see AAW48432) and fragments; and (3) a method for producing antibodies which specifically bind to a MA comprising administering to a mammal a plasmid comprising a sequence which encodes at least 1 ME derived from a polypeptide having an amino acid sequence (see AAW48432) and fragments. The products and methods can be used for detecting, diagnosing, staging, monitoring, prognosticating, preventing or treating, or determining predisposition to diseases or conditions of the breast such as breast cancer.

CC Sequence 93 AA;

Query Match 49.8%; Score 281; DB 19; Length 93;

Best Local Similarity 58.7%; Pred. No. 5.7e-23; Mismatches 25; Indels 0; Gaps 0;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

14 MKLIMVLMALALHLCYADSGCKLDEDMVEKTI NSDISPEYKELQEFIDSDAAAEAMG 73

Db 1 mklimvmlaalsqhcagsgcpllenviskrlmpgskvkeykellqelidnatnaid 60

QY 74 KFKOCFLNOSHRTLUKNFGLMHHTVYDSIWCNM 105

Db :|||||: ||| :| :||| :||
61 elkeclfnqdelisnevfmqllydsjcdl 92

RESULT 13

AAV01718
ID AAV01718 standard; Protein; 93 AA.

AC AAV01718;

DT 25-JUN-1999 (first entry)

XX Mammaglobin, a mammary specific protein.

DE Human; mammary-specific protein; mammaglobin; antigen; vaccine;

XX mammary-specific protein; mammary-specific protein; antigen; vaccine;

KW mammary-specific protein; mammary-specific protein; antigen; vaccine;

OS mammary-specific protein; mammary-specific protein; antigen; vaccine;

PN WO914230-A1.

PD 25-MAR-1999.

PF 18-SEP-1998; 98WO-US17991.

PR 18-SEP-1997; 97US-0933149.

XX (UNIW) UNIV WASHINGTON.

PI Fleming TP, Watson MA;

PI Fleming TP, Watson MA;

DR WPI; 1999-244021/20.

DR N-PSDB; AAX26966.

XX Mammaglobin, secreted protein overexpressed in breast cancer

PT Claim 15; Fig 2; 60pp; English.

The present sequence represents a human mammary-specific protein, designated mammaglobin. The specification describes a protein comprising a mammaglobin antigen that is recognized by B and/or T cells specific for the natural, secreted and glycosylated form of mammaglobin polypeptide. This protein, or recombinant vectors that express it, are used in vaccines for treating mammaglobin-expressing cancers, specifically of the breast. Such cancers can also be treated using autologous tumor lymphocytes activated ex vivo with an mammaglobin antigen, then returned to the patient. Expression of mammaglobin is elevated in 27% of stage I primary breast cancers, so it represents a marker useful for diagnosis of this disease.

CC Sequence 93 AA;

Query Match 49.8%; Score 281; DB 20; Length 93;

Best Local Similarity 58.7%; Pred. No. 5.7e-23; Mismatches 25; Indels 0; Gaps 0;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

14 MKLIMVLMALALHLCYADSGCKLDEDMVEKTI NSDISPEYKELQEFIDSDAAAEAMG 73

Db 1 mklimvmlaalsqhcagsgcpllenviskrlmpgskvkeykellqelidnatnaid 60

QY 74 KFKOCFLNOSHRTLUKNFGLMHHTVYDSIWCNM 105

Db 61 elkeclfnqdelisnevfmqllydsjcdl 92

RESULT 14

AAAB13786
ID AAAB13786 standard; Protein; 93 AA.

XX AAAB13786;

```

XX 20-JUN-2001 (first entry)
XX Human mammaglobin.
XX DE
XX Human: breast cancer; breast disease detection; mammaglobin;
XX KW uteroglobin; chromosome 11q13; B0101; endometrial; cytostatic.
XX OS
XX Homo sapiens.
XX FH
XX Key Location/Qualifiers
XX FT Modified-site 53 /note= "Optionally N-glycosylated"
XX FT Modified-site 68 /note= "Optionally N-glycosylated"
XX FT
XX W0200035950-A2.
XX PN
XX 22-JUN-2000.
XX PD
XX 20-DEC-1999; 99WO-US30489.
XX PF
XX 18-DEC-1998; 98US-0215818.
XX PR
XX (ABBO ) ABBOTT LAB.
XX PA
XX Colpits TL, Russell JE;
XX PI
XX WPI: 2000-442366/38.
XX DR N-PSDB; AAA64845.
XX PT
XX Multimeric polypeptide antigen and antibody specific to the antigen are
XX PT useful for diagnosing, detecting and treating breast cancer -
XX PS
XX Claim 1; Pages 123-124; 124pp; English.
XX CC
XX Mammaglobin is a member of the uteroglobin protein family. The
XX CC mammaglobin gene has been localised to chromosome 11q13. The present
XX CC sequence is the protein sequence for human mammaglobin. The present
XX CC invention relates to a multimeric polypeptide antigen, which comprises
XX CC of the present sequence and B0101 polypeptide (AA13787). B0101 is
XX CC another uteroglobin protein. The presence of multimeric polypeptide
XX CC antigen in a test sample can be used as the basis for a test to diagnose
XX CC breast disease e.g. breast cancer, in a patient. The detection can be
XX CC carried out using antibodies specific for the multimeric polypeptide
XX CC antigen.
XX SQ
XX Sequence 93 AA:

Query Match 49.8%; Score 281; DB 21; Length 93;
Best Local Similarity 58.7%; Pred. No. 5.7e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 14 MKLWMLMALLLHCYADSGCKLEDMWEKTNISDISPEYKELIQFIDSDAAAEAMG 73
DB 1 mklwmlmaalsqhcysgscpllenvisklnpyskleykellqellidnatnaid 60
OY 74 KFKQCFLNQSHRTLNKFGMLMHTVYDSICNM 105
DB 61 elkecfingtdetlsnvevfmgjlydssldcl 92

RESULT 15
AA184622
ID AAY84622 standard; Protein; 93 AA.
XX
XX AAY84622;
AC
XX
XX 25-JUL-2000 (first entry)
DT
XX
XX Amino acid sequence of the mammary-specific protein mammaglobin.

```

```

KW Human; mammaglobin; mammary gland; breast cancer; endometrial cancer.
XX OS
XX Homo sapiens.
XX PN W0200018783-A1.
XX PD 06-APR-2000.
XX PF
XX 29-SEP-1999; 99WO-US22616.
XX PR
XX 29-SEP-1998; 98US-0162622.
XX PA
XX (UNIV ) UNIV WASHINGTON.
XX PI
XX Watson MA, Fleming TP;
XX PN
XX WPI: 2000-293105/25.
XX DR N-PSDB; AAA12632.
XX PT
XX Methods for detecting breast cancer, comprising detecting elevated
XX PT concentrations of a mammaglobin polypeptide, using an antibody, or
XX PT detecting elevated concentrations of the mRNA encoding the polypeptide,
XX PT using oligonucleotides -
XX PS
XX Example 1; Fig 2; 71pp; English.
XX CC
XX The present sequence represents the human mammary-specific secreted
XX CC protein mammaglobin. Mammaglobin expression is restricted to the
XX CC mammary gland. Dysregulation of the mammaglobin gene occurs early
XX CC and frequently in breast cancer. The specification describes a method
XX CC for detecting the presence of breast cancer in a patient, comprising
XX CC detecting an elevated concentration of mRNA encoding a mammaglobin
XX CC polypeptide. The methods are useful for detecting the presence of
XX CC breast and endometrial cancer.
XX SQ
XX Sequence 93 AA:

Query Match 49.8%; Score 281; DB 21; Length 93;
Best Local Similarity 58.7%; Pred. No. 5.7e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 14 MKLWMLMALLLHCYADSGCKLEDMWEKTNISDISPEYKELIQFIDSDAAAEAMG 73
DB 1 mklwmlmaalsqhcysgscpllenvisklnpyskleykellqellidnatnaid 60
OY 74 KFKQCFLNQSHRTLNKFGMLMHTVYDSICNM 105
DB 61 elkecfingtdetlsnvevfmgjlydssldcl 92

Search completed: June 20, 2002, 11:02:43
Job time: 59 sec

```


RESULT 1
US-08-821-451A-6
Sequence 6, Application US/08821451A

Patent No. 6066724

GENERAL INFORMATION:

APPLICANT: Jian NI, Guo-Liang Yu and Reiner Gentz

TITLE OF INVENTION: Human Endometrial Specific Steroid-

TITLE OF INVENTION: Binding Factor I, II and III

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

ADDRESSEE: CECCHI, STEWART & OLSTEIN

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/821,451A

FILING DATE: March 21, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/014,724

FILING DATE: March 21, 1996

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-521 (PP257)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 95 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

US-08-821-451A-6

Query Match 87.9%; Score 496; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 8,8e-54;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Thu Jun 20 11:26:58 2002

us-09-673-39

QY	14	MLKLVLMALALHLCYADSGCKLLEDVYKTIINSISIPYKELLOEFIDSDAAEANG	73
DB	1	MLKLVLMALALHLCYADSGCKLLEDVYKTIINSISIPYKELLOEFIDSDAAEANG	60
QY	74	KFROCFINOSHRTLNKFGIMAHVYDSTICNKSIN	108
DB	61	KFROCFINOSHRTLNKFGIMAHVYDSTICNKSIN	95

RESULT 2

PS Claim 23; Page 314; 444pp; German.
 XX This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer; (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for
 CC generation of specific antibodies. (A) are identified by assembling ESTs
 CC (expressed sequence tags) from a particular tissue type before comparison
 CC of expression patterns. This allows a significantly longer fragment of
 CC the gene to be revealed, so should reduce the number of failures
 CC associated with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. AAY5941-Y60328 represent
 CC protein fragments encoded by the human endometrium tumour cDNA library
 CC derived EST fragments represented in AAY41901-242121.

SO Sequence 108 AA:

Query Match 100.0%; Score 564; DB 20; Length 108;
 Best Local Similarity 100.0%; Pred. No. 1.5e-53;
 Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CHARLNTDSSRLAMKLLMVLMTALLHCYADSGCKLLEDVMEKTIINSIDISPEYKELIQ 60
 DB 1 charlntdsrrlamlklmvlmtallaallhcyadsgckllledvmeaktinsidsispeykellq 60
 QY 61 EFIDSDAAAEAMGKRFKQCFINOSHRILKNGLAMHTVYDSIMCNMKS 108
 DB 61 efidsdaaaeamgkfkqcfinshrilknnglammhtvydsiwcnmksn 108

RESULT 2

AAW35804 standard; Protein; 95 AA.

AC AAW35804;

DT 27-MAR-1998 (first entry)

XX Human endometrial specific steroid-binding factor III.

KW Endometrial specific steroid-binding factor III; ESF III; human;

KW Clara cell secretory protein; endometrium;

KW phospholipase A2 inhibitor; polychlorinated biphenyl; antiagregant;

KW inflammation; asthma; rhinitis; cystic fibrosis; airway disease;

OS neoplasia; atopy; therapy; diagnosis.

OS Homo sapiens.

FT Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT Protein 22..95
 FT /label= Mat_protein

XX WO9734997-A1.

XX 25-SEP-1997.

XX 21-MAR-1996; 96WO-US03857.

XX 21-MAR-1996; 96WO-US03857.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Gentz RL, Ni J, Yu G;

WPI: 1997-480206/44.

NCBI: AAT94832.

XX Human endometrial specific steroid-binding factor I, II and III -
 PT used to treat inflammation, asthma, rhinitis, cystic fibrosis,
 CC airway disease, neoplasia, atopy etc.

PS Claim 19; Page 65; 92pp; English.

XX This sequence comprises human endometrial specific steroid binding
 CC factor III (ESF III), a protein that inhibits phospholipase A2
 CC activity, binds to polychlorinated biphenyl compounds, reduces
 CC foreign protein antigenicity, inhibits monocyte and neutrophil
 CC chemotaxis and phagocytosis, inhibits platelet aggregation,
 CC regulates eicosanoid levels in the human uterus and controls the
 CC growth of endometrial cells. The amino acid sequence was deduced
 CC from a cDNA clone (see AAT94832) derived from a human endometrial
 CC tumour. ESF I (see AAW35802) and ESF II (see AAW35803) are also
 CC claimed. Human ESF III has about 36% identity with rat prostatic
 CC steroid-binding protein. Recombinant ESF I, II and III can be
 CC expressed in host cells for use in claimed methods (a) for treating
 CC a patient in need of ESF I, II or III (including expression of the
 CC polypeptide in vivo) and (b) for identifying compounds which bind
 CC to and inhibit activation of the ESF polypeptide. ESF I, II and
 CC III may be used to treat inflammation, asthma, rhinitis, cystic
 CC fibrosis, airway disease, neoplasia and atopy.

SO Sequence 95 AA:

Query Match 87.9%; Score 496; DB 18; Length 95;
 Best Local Similarity 100.0%; Pred. No. 2.9e-46;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 MKLMTVMTLAAALHLCYADSGCKLLEDVMEKTIINSIDISPEYKELIQEFIDSDAAAEAMG 73
 DB 1 mklmvtmllaalhlhcyadsgckllledvmeaktinsidsispeykellqefidsdaaaeamg 73

QY 74 KFKOCFLINOSHRILKNGLAMHTVYDSIMCNMKS 108
 DB 61 kfkqcfinshrilknnglammhtvydsiwcnmksn 95

RESULT 3

AAW02590 standard; Protein; 95 AA.

AC AAW02590;

DT 26-JUL-1999 (first entry)

XX A human mamoglobin homologue (HMH).

KW Human mamoglobin homologue; HMH; antagonist; neoplastic disorder;

KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;

KW teratocarcinoma; endometriositis.

OS Homo sapiens.

PN WO9919487-A1.

XX 22-APR-1999.

XX 14-OCT-1998; 98WO-US21729.

XX 16-OCT-1997; 97US-0951750.

XX (INCY-) INCYTE PHARM INC.

XX Hillman JL, Murry LE, Shah P;

DR WPI: 1999-302531/25.

DR N-PSDB; AAX36138.

PT New human mamoglobin homologue (HMH), useful for diagnosing, treating